

From: Jiang, Dong
Sent: Tuesday, April 26, 2005 11:48 AM
To: STIC-Biotech/ChemLib
Subject: 10/075,846

Please search 1) SEQ ID NO:3

2) nucleotides 574-621 of SEQ ID NO:3

-pending
-issued & Pub.
-commercial

Please send results on paper to Dong Jiang in REM 4D70 (mail stop REM 4C70).
Thank you very much.

Dong Jiang

AU1646
REM - 4D70
571-272-0872
Mail stop REM - 4C70

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 4/27/05
Date Completed: 4/29/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 2 AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: ot
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

1	1521.2	92.8	1874	6	AX392950
2	1503.8	91.7	1600	6	AX037565 Sequence
3	1002.2	61.1	1251	10	AF462147 Mus musculus
4	695.8	42.4	2045	5	DRE040970 Danio rerio
5	659.2	40.2	1384	10	RNGRSTR R.norvegicus
6	656.2	40.0	1407	10	S73717 GIRAL=Inhibitor
7	656	40.0	1439	9	BC074980 Homo sapiens
8	656	40.0	1712	6	CQ278075 Sequence
9	656	40.0	1715	9	HSLGLYRA2 H.sapiens a
10	654.8	39.9	1359	10	RNO310837 Rattus norvegicus
11	654.8	39.9	3160	10	RNIGPAA2 R.norvegicus
12	654.8	39.9	3865	6	AX700251 Sequence
13	654.8	39.9	3865	10	RNNEOGLY Rat NGlyc m
14	654.6	39.9	1350	10	RNO310834 Rattus norvegicus
15	654.6	39.9	1388	10	AY129229 Mus musculus
16	654.6	39.9	2125	10	RATGILYRAL Rattus norvegicus
17	653.2	39.8	2775	10	BC056342 Mus musculus
18	653.2	39.8	2775	10	BC068987 Mus musculus
19	650.2	39.6	1857	9	HSLGLYRA1 H.sapiens a

Qy	1	ATGACAACTCTGTTGCTCCTGCAACCCCTCTCCTTCTCTTCTCTGGAACCTTCGCAGGCGAG	60
Db	286	ATGACAACTCTGTTGCTGCAACCCCTCTCCTTCTCTTCTCTGGAACCTTCGCAGGCGAG	345
Qy	61	GTCTCTCCTCAGGGTGCCCTTGGCAAAAGAGGAAGTCAAAATCTGGAAACCAAGGGGTCCCG	120
Db	346	GTCTCTCCTCAGGGTGCCCTTGGCAAAAGAGGAAGTCAAAATCTGGAAACCAAGGGGTCCCG	405
Qy	121	CCCATCTGCCCCCTCTGATTTCTTAGACAAACTTATGGGGCGAAACATCTGGATATGATGCC	180

	D	b	1	A	T	G	C	A	A	C	T	T	G	T	T	C	T	G	C	A	A	C	C	T	C	T	C	T	T	C	T	T	C	T	T	C	T	T	C	T	T	C	T	T	C	T	G	G	A	C	C	T	T	G	C	C	A	G	G	C	A	G	60
	Q	y	61	G	T	C	T	C	T	C	A	G	G	T	G	G	C	T	T	G	C	A	A	A	A	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	120			
	D	b	61	G	T	C	T	C	T	C	A	G	G	T	G	G	C	T	T	G	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	120				
	Q	y	121	C	C	A	T	G	T	C	C	C	C	T	C	T	G	A	T	T	C	T	A	G	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	180				
	D	b	121	C	C	A	T	G	T	C	C	C	C	T	C	T	G	A	T	T	C	T	A	G	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	180				
	Q	y	181	A	G	A	T	T	C	G	G	C	C	A	A	T	T	T	A	A	A	G	C	C	A	C	C	C	G	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	T	T	C	A	A	C	A	A	C	A	240			
	D	b	181	A	G	A	T	T	C	G	G	C	C	A	A	T	T	T	A	A	A	G	C	C	A	C	C	C	G	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	T	T	C	A	A	C	A	A	C	A	240			
	Q	y	241	A	G	T	T	C	A	G	T	C	C	G	T	C	A	C	A	A	G	A	C	C	A	C	A	A	A	G	A	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	300				
	D	b	241	A	G	T	T	C	A	G	T	C	C	G	T	C	A	C	A	A	G	A	C	C	A	C	A	A	A	G	A	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	300				
	Q	y	301	C	A	G	T	G	A	A	T	G	A	C	C	G	T	G	T	C	T	A	C	G	A	A	A	T	A	T	C	T	C	T	G	A	N	A	C	T	C	T	G	A	N	A	C	T	C	T	G	A	N	A	C	T	C	G	A	360			
	D	b	301	C	A	G	T	G	A	A	T	G	A	C	C	G	T	G	T	C	T	A	C	G	A	A	A	T	A	T	C	T	C	T	G	A	N	A	C	T	C	T	G	A	N	A	C	T	C	T	G	A	N	A	C	T	C	G	A	360			
	Q	y	361	C	C	T	C	C	A	T	G	T	G	A	C	T	A	T	C	T	G	A	A	G	C	C	A	C	C	T	T	T	T	G	C	T	A	N	A	G	A	A	G	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	420				
	D	b	361	C	C	T	C	C	A	T	G	T	G	A	C	T	A	T	C	T	G	A	A	G	C	C	A	C	C	T	T	T	T	T	G	C	T	A	N	A	G	A	A	G	G	G	C	C	C	C	C	C	C	C	C	C	C	C	420				
	Q	y	421	A</																																																											

Qy	1141	CAAGAAAGTCGTTTCTATATTTCCGTTGGGCTATGGCTTTGGGCCACTCTCTGCAGGCAAGAGAT	1200
Db	1099	CAAGAAAGTCGTTTCTATATTTCCGTTGGGCTATGGCTTTGGGCCACTCTCTGCAGGCAAGAGAT	1158
Qy	1201	GGAGGTCCAATGGAAGGTTCTGCGATTTATAGTCCCAACTCCAGGCCCTCTTCTTAAGG	1260
Db	1159	GGAGGTCCAATGGAAGGTTCTGCGATTTATAGTCCCAACTCCAGGCCCTCTTCTTAAGG	1218
Qy	1261	GAAGGAGAAACCAACGCGGAAACTCTACGTGGACTGAGCCAAAGAGAAATTGACACCATCTCC	1320
Db	1219	GAAGGAGAAACCAACGCGGAAACTCTACGTGGACTGAGCCAAAGAGAAATTGACACCATCTCC	1278
Qy	1321	C-GGGGCTGCTTCCGCTTTCACTTTCCATCTTCATATCTTCAATCTCTACCTGGGTTGTCTATAA	1379
Db	1279	CAAGGGCGAAATCCCTTTCACTTTCCATCTTCAATCTCTTCAATCTCTTCTACTGGGTTGTCTATAA	1338
Qy	1380	AGTGCTATGCTCAGAAGATATCCACCAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCC	1439
Db	1339	AGTGCTACGGTCAGAAGATATCCACCAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCC	1398
Qy	1440	TGCTGTGGGCTCCTGCTTCTCTCTGGTGGGCTTTCTCCCTCAGTTPAGACTCCATPAGG	1499
Db	1399	TGCTGTGGGCTCCTGCTTCTCTCTGGTGGGCTTTCTCCCTCAGTTPAGACTCCATPAGG	1457
Qy	1500	GGTTTGACAGTTCCTTCTTGATCTCCCACTCAGAACTTCAACTACCAGTCCCAAGACTA	1559
Db	1458	GGTTTGACAGTTCCTTCTTGATCTCCCACTCAGAACTTCAACTACCAGTCCCAAGACTA	1517
Qy	1560	TGTGGGCTATATTCGATGCTGCAATGGTGGCTGTACTATTATAAGATGGCTTATCTACC	1619
Db	1518	TGTGGGCTATATTCGATGCTGCAATGGTGGCTGTACTATTATAAGATGGCTTATCTACC	1577
Qy	1620	CTAAAAAATAAAAAAAAAAAAAA 1640	
Db	1578	CCTAAAAAATAAAAAAAAAAAAAA 1598	
RESULT 3			
AF462147			
LOCUS			
DEFINITION	AF462147	1251 bp	mRNA linear ROD 11-FEB-2002
			Mus musculus glycine receptor alpha 4 subunit (Glr4) mRNA,
			complete cds.
ACCESSION	AF462147		
VERSION	AF462147.1	GI:18448710	
KEYWORDS			
SOURCE			
ORGANISM			
			Mus musculus (house mouse)
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE			
AUTHORS			
TITLE			
			Localization of different glycine receptor isoforms in murine
			spinal cord
JOURNAL			
REFERENCE			
			Unpublished
			2 (bases 1 to 1251)
AUTHORS			
TITLE			
			Groemer,T.-W., Becker,C.-M. and Becker,K.
JOURNAL			
			Direct Submission
			Submitted (21-DEC-2001) Biochemistry, University of Erlangen,
			Fahrstrasse 17, Erlangen 91054, Germany
FEATURES			
source			
			1. .1251
			/organism="Mus musculus"
			/mol_type="mRNA"
			/strain="B6C3/Fe"
			/db_xref="taxon:10090"
			/chromosome="X"
			/tissue_type="spinal cord"
gene			
			1. .1251
			/gene="Glr4"
			/note="synonym: mmGlr4"
CDS			
			1. .1251
			/gene="Glr4"

/clone="GR1"	
/tissue type="spinal cord"	
<1..1286	
/codon start=3	
/product="glycine receptor strychnine binding subunit"	
/protein_id="CAA68378.1"	
/db_xref="GI:755778"	
/db_xref="GOA:P07727"	
/translation="UniProt/Swiss-Prot:P07727"	
/tranelation="SKEVDAARSAPKSPSPDLKLMGRTSGYDARIPNFKGPPVN	
VSCNFINFSGIAETMDYRNIPLRQNDPRLAYNEYPDSDLDLSLWSKP	
DUFFANEKGAHFEITDKNGKLSRGNVLYSIRITLLACPMDLKNPMDVQICIM	
QLESFGYNDLIFEFQEGCAQVADGLTPQILKBEKDLRYCTKHYNTKFTCEIA	
RFLHRQMYIILIPISLLVILSWIFWINMDAAPARVGLGITVLTMTTQSSGS	
RASLPKVSIVLIDIMWCLLFVFSALLEVAANFVSRQHKELLRLFKTRHHKDE	
GEGGRNFSAWGMGPACLOAKDGISVKGANNTNPAPAPSKSPPEMRKLIQRAKK	
IDKLSRIGFPMALFIPNMFYIYKIVREDVHNK"	
21..1283	
mat_peptide	/product="glycine receptor strychnine binding subunit"
misc_feature	561
/note="c/t transition"	
ORIGIN	
Query Match 40.2%; Score 659.2; DB 10; Length 1384;	
Best Local Similarity 72.1%; Pred. No. 6.3e-170;	
Matches 954; Conservative 0; Mismatches 313; Indels 57; Gaps 5;	
QY	89 AGGAAGTCAAAATCGGAACCMAGGGTCCCGCCCATGTCCTCGTGAATTCCTAGACA 148
Db	7 AGAGGTGACGCTGCGCGCTCGACCCCAAGGCTATGACCCCTCGGACTTCCTGGATA 66
QY	149 AACTATGGGGCGAACAATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC 208
Db	67 AGCTATGGGAGGACTTCTGGGTATGATGCCAGGATCAGACCACTTTAAAGGTCCTC 126
QY	209 CGGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGTCCGTCACAAAGCACAA 268
Db	127 CTGTGAACGTGAGTTGCAACATCTTCATCAACAGCTTTGGTTCTATCGCGGAGACAAC 186
QY	269 TGGACTACCGGGTGAATGCTCTTGGCGCACAGTGGATGACCCACCGCTGCTTACC 328
Db	187 TGGACTACAGGGTCAACATCTTCTGAGGAGCAGTGGAAACGACCCCGCTGCGCCTACA 246
QY	329 GAGAATATCTGATGACTCTCTGGACTCGATCCCTCCATGCTGGACTCTATCTGGAAGC 388
Db	247 ATGAATATCTGACGACTCTGGAACCTTGACCCATCAATGTTGATTCATCTGGAAGC 306
QY	389 CAGACCTCTTTTGTGTAATGAGAAAGGGCCCACTTCCATGAGGTGACCAAGCAACA 448
Db	307 CTGACTTGTTTGTGCAATGAGAGGGGGCCCACTTCCACGAGATCACCACGGAACA 366
QY	449 AGTTACTGCGCATCTTCAAGATGGGAATGCTGTGTACAGCATCAGGCTGACCCCTCATTT 508
Db	367 AGCTGCTGAGAATCTCCCGAAGCGCAAGCTCTCTACAGCATCAGATCACCCTGACTC 426
QY	509 TGTCTGCTGATGAGACCTCAAGAACTTCCCATGACATCCAGAGTGCACGATCAGC 568
Db	427 TGGCTTCCCCCATGGACCTGAAGATTTCCGATGACGCTACAGACATGATATGCAAC 486
QY	569 TTGAGAGCTATCCATFACTCTGAGCGCCTTGCCACTCTGTGACATTTTCAGTTGGGTACA 628
Db	487 TGGAAAGCT-----ITGGTTATA 504
QY	629 CCATGAAGACCTCGTGTGTGAGTGTGGAAGATGCTCTCTGCTGCCAAGTGGCTGAGG 688
Db	505 CCATGAAGACCTCATCTTTTGAATGTC---AAGCAAGAGCTGTCAGGTGGCAGATG 561
QY	689 GGCTGACTCTGCCCGCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTGTGACCAAGC 748
Db	562 GACTGACCTGCTCAGTTTATCTTGAAGAGAGAGGATCTGAGATCTGACACTGCACCAAGC 621
QY	749 ACTAACAACAGGGAATTCACCTGATCGAGTAAAGTTTACCTGGAAACGGCAGATGG 808

Db	622 ACTACAACACAGGTAAATTCACCTGCATTTGAGGCCGCGATTCCACCTGGAAACGGCAGATGG 681
QY	809 GCTACTATCTGATTCAGATGTACATCCCGAGCTTACTATCGTCTATCTCTGCTCTGGGTCT 868
Db	682 GCTACTATCTGATTCAGATGTACATCCCGAGCTTACTATCGTCTATCTCTGCTCTGGATCT 741
QY	869 CTTTCTGGATCAACATGATGCTGCCCCCTGCCCCGCTGTGGCTGGGCTGGGCATCAACACCGTGC 928
Db	742 CTTTCTGGATCAACATGATGCTGCCCCGCTGTGGCTGGGCTGGGCATCAACACAGTGC 801
QY	929 TCACATGACACACAGAGCTTGGCTCCCGGCTCTTTGGCTAAGGTGTCTTACGTCA 988
Db	802 TCACATGACACACAGAGCTTGGCTCCCGAGCTTGGCTCCCGAGGTGTCTTACGTCA 861
QY	989 AGGCAATCAGATCTGGATGGCTGTGTCTCTTTGTGTGTCTGCTGCTGCTGCTGCTGCTGCT 1048
Db	862 AGCTATTGACATCTGGATGGCTGTGTCTCTTTGTGTGTCTGCTGCTGCTGCTGCTGCTGCT 921
QY	1049 ATGCTGCCATAAATTTTCTCTGTCAGCATPAAAGAAATTCATACGACTTCGAAGAAGGC 1108
Db	922 ATGCCGCTGTCAACTTTGTGTCGCCGCAACACAGGAATC-----CTTCGATTTAGGA 975
QY	1109 AGAGGGCCCAACGCTTGGAGGAGATATCATCCAGAAAGTCTGTTCTATTTCCGTGGCT 1168
Db	976 GGAAGCGCGCATCATCAAGGATGATGAGGGTGGGAGAGGCCGCTTCAACTTCTCCGCT 1035
QY	1169 ATGGCTTGGGCC---ACTGCTGAGGCAAGAGATGGAGGTCCAATGGAAGTCTTGGA 1225
Db	1036 ATGGATGGGCCAGCTGCTGCAAGCAAGATGGCATCTCTGTCAAGGTGGCAACA 1095
QY	1226 TTTATAGTCCCAACCTCCAGCCCT---CTTCTAAGGGAAGAGAAACACCGCGGAAC 1282
Db	1096 ACAACAACACACGAAACCCGCTCTGCAACCGTCCCAAGTCCCGGAGGAGATGCGGAAC 1155
QY	1283 TCTAGTGAGCTGAGCAAGCAAGAAATGACACATCTCCGGGCTGTCTCCCTTTCACTT 1342
Db	1156 TCTTCATCAGAGAGCCCAAGAAATATCGCATCGGTTTTCCTCCCTGGCT 1215
QY	1343 TCCTCATCTTCAATCTTCTACTGCTGCTGCTATAAAGTGTATGCTCAGAGATATCC 1402
Db	1216 TCCTCATCTTCAATCTTCTACTGATCTACAGATCTCGGAGAGAGGAGCTCC 1275
QY	1403 ACCA 1406
Db	1276 ACA 1279

RESULT 6	
S73717	
LOCUS	
DEFINITION	
S73717	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
REMARK	
FEATURES	
source	
1..1407	
/organism="Mus sp."	
/mol_type="mRNA"	
S73717	
1407 bp	
mRNA	
linear	
ROD 17-JUL-2000	
GLR1=Inhibitory glycine receptor alpha 1 subunit [mice, A/HeJ,	
brainstem, spinal cord, mRNA, 1407 nt].	
S73717	
S73717.1	
GI:765206	
Mus sp.	
Mus sp.	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 (bases 1 to 1407)	
Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L.,	
Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P. and O'Connell, P.	
A missense mutation in the gene encoding the alpha 1 subunit of the	
inhibitory glycine receptor in the spasmodic mouse	
Nat. Genet. 7 (2), 131-135 (1994)	
95004575	
7920629	
GenBank staff at the National Library of Medicine created this	
entry [NCBI gibbs 156492] from the original journal article.	
Location/Qualifiers	
1..1407	
/organism="Mus sp."	
/mol_type="mRNA"	

RESULT 7	
BC074980	
LOCUS	1439 bp mRNA linear PRI 04-AUG-2004
DEFINITION	Homo sapiens glycine receptor, alpha 1 (startle disease/hyperplexia, sciff man syndrome), mRNA (cDNA clone MGC:103909 IMAGE:30915298), complete cds.
ACCESSION	BC074980
VERSION	BC074980.2 GI:50959672
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1439) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeng,H., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,C.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villaon,D.K., Huzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Query Match	40.0%	Score 656.2	DB 10	Length 1407
Best Local Similarity	72.6%	Pred. No. 4.3e-169		
Matches 942	Conservative 0	Mismatches 298	Indels 57	Gaps 5
QY	116	CCGAGCCCATGTCCTCTGATTCTTCTAGACAAACTTATGGGCGGAACATCTGGGATG	175	
DB				
QY	155	CCAAGCCTATGTACCCCTCGGACTTCTGTGAAGCTCATGGGAGGACTTCTGGGTATG	214	
DB				
QY	176	ATGCCAGGATTTCGGCCCAATTTTAAAGGCCCCACCGGTGAACGTGACCTGCAACATCTTCA	235	
DB				
QY	215	ACGCCAGGATCAGGCCCACTTTAAGGTCTCTGTGAACGTAAAGTTGGCAACATCTTCA	274	
DB				
QY	236	TCAACAGTTTCAGCTCCGTCCACCAAGACCAATGGACTACCGGGTGAATGTCTTCTTGC	295	
DB				
QY	275	TCAACAGTTTCGGTTCATCGCTGAGACAACCATGGACTATAGGTCACACATCTTCTCTGA	334	
DB				
QY	296	GGCAACAGTGGGATGACCCAGCGCTGTCTTACCGAGNAATATCTGTATGACTCTCTGGACC	355	
DB				
QY	335	GGCAGCAGTGGGAACGACCCCGCTGTGCGCTTACAAATGAATACCTGTATGACTCTCTGGACC	394	
DB				
QY	356	TGCATCCTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTCTTTTGTCTAATGAGAAAG	415	
DB				
QY	395	TTGACCACTATGTTGGATTCCATCTGGAAGCTGACTTGTCTTTGGCAATGAGAAGG	454	
DB				
QY	416	GGGCCCAACTTCCATGAGGTGACCAACGAGCAACAAAGTTTACTCGCGCATCTTCAAGAAATGGGA	475	
DB				
QY	455	GGSCCCACTTCCACGAAATCACCACGAGCAACAACTGCTAAGAACTCCCGGAATGGCA	514	
DB				
QY	476	ATGTGCTGTACAGATCAGGCTGACCTCTCATTTTGTCTCGCTGTATGGAACCTCAAGAACT	535	
DB				
QY	515	ATGTCTCTACAGCATCAGAATCACCTGACGCTGCGCTGCCCATGGACCTCAAGAAAT	574	
DB				
QY	536	TCCCATGGACATCCAGACGTGACAGTGCAGCTGTGAGACTCATCATACTCTGCAGCC	595	
DB				
QY	575	TCCCGATGGATGTACAGACGTGTATCATGCAATCGAAAGCT	616	
DB				
QY	596	CTCTGCCATCTCTGTCACTTTTCAGTTTGGCTTACACCATGAAAGACCTCGTGTTTGAGTGGC	655	
DB				
QY	617	-----TTGGATATACCATGAACGACCTCATCTTTGAGTGGC	652	
DB				
QY	656	TGGAAGATGCTCCTGTGTCCTCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTTCG	715	
DB				
QY	653	AGGAGCAAG--GAGCTGTGCAGGTGSCAGACGACTGACCTCTGCCTCAGTTTATTCTGA	709	
DB				
QY	716	GGGATGAGAAGGATTAGGCTGTTGTATCCAAAGCACTACAAACAGGGAATATCACCTGCA	775	
DB				
QY	710	AGGAAGAGAAGACCTGAGATCTGCAACCAAGCACTACAAACAGGTAAATATCACCTGCA	769	
DB				

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1439)
 Director MGC Project.
 Direct Submission
 Submitted (25-JUN-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgi.nci.nih.gov>
 On Aug 4, 2004 this sequence version replaced gi:49901657.
 Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Genome Sequence Centre, British Columbia Cancer
 Center
 cDNA Library Preparation: British Columbia Cancer Research Center
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
 Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu,
 Parvaneh Saeedi, JR Santos, Angeli Schnerch, Ursula Skalska,
 Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacquie
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRBU Plate: 2 Row: C Column: 11.
 Location/Qualifiers

1..1439
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:103909 IMAGE:30915298"
 /tissue_type="Pooled tissue, PCR rescued clones"
 /clone_lib="NIH MGC_245"
 /notes="Vector: pCR4 Topo TA"
 1..1439
 /genes="GLR1"
 /notes="synonym: STHE"
 /db_xref="LocusID:2741"
 /db_xref="MIM:138491"
 42..1391
 /genes="GLR1"
 /product="glycine receptor, alpha 1"
 /protein_id="AAH74980.1"
 /db_xref="GI:49901658"
 /db_xref="LocusID:2741"
 /db_xref="MIM:138491"
 /translation="MYSPNTLRLLYMETIVFFSLAASKEAARSAPKPMSPDFLDK
 LMGRTSGYDARPNPKGPPVNVSCNIFINSFGSAETITMDYRVNIFLRQWNPRLA
 VNEYPDSDLDPSMDSIWKPLDFANEKGAHFHEITTDNKLRLSRNGVLYSIRI
 TLTLACPMDLKQFMDVQTCIMQLESFGYTMNDLIFEMQEGAVOVADGLTLPQITLK
 EKDLRYCTKHYNGFTICIEARFLHERONGYLYIOMYPSLLIVLISWISWINDA
 APARVLGTLVTMTTSSGRASLPYKVIDIWMVCLLFVFSALLEYAVNF
 VSRQKELLRRKRHHKDEKAGEKFNFSAYMGPCALQAKDGLSVKGNNSNTN
 PPPAPSKSPKMRKLFQRAKKIDKISRGFMFLFNWYIIVKIVRRDVRHQ"

Query Match 40.0%; Score 656; DB 9; Length 1439;
 Best Local Similarity 71.9%; Pred. No. 4.8e-169;
 Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;

89 AGGAAGTCAAATCTGTGAACCAAGGGGTCCAGCCCATGTCCCTCTGTATTTCTTAGACA 148
 112 AGGAGCTGAAGTCTGTCTCGCTCCGACCAAGCTTATGTACCTCGATTTCTTGGATA 171
 149 AACTTATGGGGGAAACATCTGTGATATGATGCGAGATTCGGCCCAATTTAAAGGCCAC 208
 172 AGCTAATGGGAGAACCTCCGGATATGATGCGAGGATCAGGCCCAATTTAAAGGTCCCC 231
 209 CGGTGAAGTGAACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCA 268
 232 CAGTGAAGTGAAGTGAACATTTTCATCAACAGCTTTGGTTCCATTTGCTGAGACCA 291
 269 TGGACTACCGGTGAATGTCTTCTGGGCAACAGTGAATGACCCACCCCTGTCTCTACC 328
 292 TGGACTATAGGTCAACATCTTCTGGGCAACAGTGAATGACCCACCCCTGTCTCTACC 351
 329 GAGAATATCTGATGATCTCTGGACCTGATGATCTCTCAATGCTGAGTCTATCTGGAAG 388
 352 ATGAATACCTGACGACTCTCTGGACCTGGACCCATCCATGCTGGACTCCATCTGGAAC 411
 389 CAGACCTCTCTCTTCTGATGAGAAAGGGGCAACTTCCATGAGTGAACGACCAACCA 448
 412 CTGACTGTCTTCTGCAACGAGAAAGGGGCAACTTCCATGAGTGAACGACCAACCA 471
 449 AGTTACTCGGACTCTTCAAGAAATGGGAATGTCTGTACAGCATCAGGCTGACCCCTCAT 508
 472 AATTGTAAGGATCTCCCGGATGGGAATGTCTCTACAGCATCAGATCACCCTGACAC 531
 509 TGTCTGCTGATGAGACTTCAAGAACTTCCCATGAGACATCCAGAGTGCAGATGACAG 568
 532 TGGCTGCCCCATGAGCTTGAAGAAATTTCCCATGAGATGTCAGACATGATATGCAAC 591
 569 TTGAGAGCTCATCCATCTCTGACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
 592 TGGAAAGCT-----TTGGATATA 609
 629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTGCTGTCTCAAGTGGCTGAG 688
 610 CGATGATGACCTCATCTTTGAGTGGCAGAACAGG---GAGCCGTGACAGGTAGCAGATG 666
 689 GGCTGACTGCTGCCAGATTTATCTTGGGGATGAGAGGATCAGGATCTAGGCTGTTGTACCAAG 748
 667 GACTAACTCTGCCCACTTATCTTGAAGGAGAGAGAGGACTTGAATATCTGCACCAAGC 726
 749 ACTACAACACAGGAAATTCACCTGCATCGAGTAAAGTTTCCCTGGAACGACAGATG 808
 727 ACTACAACACAGGAAATTCACCTGCATGAGGCGCGGTTCACCTGGAAGGAGATG 786
 809 GCTACTATCTGATTCAGATGTATCATCCCGAGCTTACTCATCTGCTCATCTCTGCTGCT 868
 787 GTTACTACCTGATCAGATGTATATCCAGGCTGTCTCATCTGCTCATCTCTCATGATCT 846
 869 CTTCTGGATTAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
 847 CTTCTGGATTAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
 929 TCACCATGACCAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
 907 TCACCATGACCAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
 989 AGGCAATCAGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
 967 AAGCCATTGATTTGGATGGCAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
 1049 ATGCTGCCATAAATTTGTTCTCGTCAAGATAAAGAAATTCATACGATTCGAGAGAGGC 1108
 1027 ATGCTGCCATAAATTTGTTCTCGTCAAGATAAAGAGAGTGTCTGCTGCTGCTGCTGCTGCTG 1086
 1109 AGAGGCGCAACGCTTGGAGGAGATATCATCAAGAAAGTGTCTTCTTCTTCTTCTTCTTCTT 1168
 1087 GGAGATCACAAGGAGGATGAAGCT-----GGAGAGGCGCTTAACTTCTTCTGCTT 1140

QY 1169 ATGGCTTGGGCC--ACTGCTGCAGGCAAGAGATGAGGTGCAATGGAAGTTC---TG 1222
 DB 1141 ATGGGATGGGCCAGGCTGTCTACAGGCCAAGGATGGCATCTCAGTCAAGGGGCCCAACA 1200
 QY 1223 GCATTATATAGTCCCAACCTCCAGCCCTCTTCTTAAGGGAAGAGAAACCCAGCGGAAAC 1282
 DB 1201 ACAGTAACACCAACACCCCTCTCTGCACCATCTAAGTCCCCAGAGGAGATGCGAAAC 1260
 QY 1283 TCTACGTGGACTGAGCGAAGAAATTTGACACATCTCCCGGGGTGTCTTCCCTTTCACTT 1342
 DB 1261 TCTTCATCAGAGGGCGCAAGATGCAAAAATATCCCGCATTTGGCTTCCCATGGCCT 1320
 QY 1343 TCTCATCTTCAATATCTTCTACTGCTGTCTATAAAGTGTATGTCAGAGATATCC 1402
 DB 1321 TCTCATTTTCAACATGTTCTACTGGATCATACAGATTGTCCTAGAGAGAGCTCC 1380
 QY 1403 ACCA 1406
 DB 1381 ACAA 1384

RESULT 8
 LOCUS CO728075 1712 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 14009 from Patent WO02068579.
 ACCESSION CO728075
 VERSION CO728075.1 GI:42295931
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
 JOURNAL Patent: WO 02068579-A 14009 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source Location/Qualifiers
 1..1712
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 40.0%; Score 656; DB 6; Length 1712;
 Best Local Similarity 71.9%; Pred. No. 5e-169;
 Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
 QY 89 AGGAAGTCAAAATCGGAACCAAGGGTCCAGCCCATGTCCCTCTGATTTCTCTAGACA 148
 DB 366 AGGAGGCTGAAGCTGTCTGCTCGCAGCCCAAGCCTATGTCAACCTCGGATTTCTTGATA 425
 QY 149 AACTTATGGGGCGAATCTGGATATGATGTCAGGATTCGGCCCAATTTTAAAGGCCAC 208
 DB 426 AGCTAATGGGGAGAACCTCCGGATATGATGCCAGGATTCAGGCCCAATTTTAAAGGTCCTC 485
 QY 209 CCGTGAACGTGACCTCAACATCTTCAACAGTTTTCAGTCCGTCCGCAACAGCACAA 268
 DB 486 CAGTGAACGTGAGCTGCAACATTTTCAACAGCTTTGGTTCATGTCGAGAACCA 545
 QY 269 TGGACTACCGGGTGAATGTCTTCTGGGCAACAGTGAATGACCCAGCCTGCTCTTACC 328
 DB 546 TGGACTATAGGCTCAACATCTTCTGGGCGAGCAATGGAACGACCCCGCTGCTATA 605
 QY 329 GAGAATATCTGATGACTCTTGGAATCTGATCCCTCCATGCTGAGTCTATCTGGAAGC 388
 DB 606 ATGAATACCTTGACGACTCTCTGGACCTGGACCCATCTCCTGAGTCTCCATCTGGAAC 665
 QY 389 CAGACCTCTTCTTTGTGTAATGAGAAAGGGCCCAACTTCCATGAGGTGACCCAGCAGCAACA 448
 DB 666 CTGACCTGTTCTTTGGCAACGAGAAGGGGGCCCACTTCCATGAGATCACCACAGCAACA 725

RESULT 9
 HSGLYRA2
 LOCUS

1715 bp mRNA linear PRI 28-MAY-1993

QY 449 AGTTACTGCGCATCTTCAAGATGGGAATGTCTGTACAGCATCAGGCTGACCCCTCATTT 508
 DB 726 AATTGCTAAGGATCTCCCGGAATGGGAATGTCTCTACAGCATCAGAATCAACCTGACAC 785
 QY 509 TGTCTGCTGCTGATGACACCTCAAGAACTCCCATGGACATCCAGAGCTGCACGATCAGC 568
 DB 786 TGGCTGCCCCATGACTTGAAGATTTCCCATGAGTGTCCAGACATGATCATGCAAC 845
 QY 569 TTGAGAGCTCATCCATATCTCTGAGCCCTCTGCCATCTCTGTACATTTTCAAGTTGGGTACA 628
 DB 846 TGGAAAGCT-----TTGGATATA 863
 QY 629 CCATGAAGACCTCTGTTTGAAGTGGTGGGAAGATGCTCCTGCTGTCCAGTGGCTGAGG 688
 DB 864 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCCGTGCGAGGTAGCAGATG 920
 QY 689 GGTGACTCTGCCCCCAGTTTATCTTTCGGGATGAGAAGGATCTAGGCTGTGTGACCAAGC 748
 DB 921 GACTAATCTCTGCCCCAGTTTATCTTGAAGGAAGAGAGGACTTGAATGACTGACCAAGC 980
 QY 749 ACTACAACACAGGGAATTCACCTGATCGAGGTAAAGTTTCACTGGAAACGCGAGATGG 808
 DB 981 ACTACAACACAGGTAAATTCACCTGATTGAGGCCCGGTTCCACCTGGAGCGGACAGATGG 1040
 QY 809 GCTACTATCTGATTCAGATGTACATCCCGAGGCTTACTCATGTCATCTCTGCTGCTGCT 868
 DB 1041 GTTACTACCTGATTCAGATGTATTTCCAGGCTGCTCATTTGTCATCTCTCATGGATCT 1100
 QY 869 CTTCTGGATCAACATGGAATGCTGCCCTGCGCTGTGGGCTTGGGCATCACCAACGCTGC 928
 DB 1101 CTTCTGGATCAACATGGAATGCTGCCCTGCTGTGGGCTTGGGCATCACCAACGCTGTC 1160
 QY 929 TCACCATGACCAACGAGCTCTGGCTCCCGGCTCTTTTGGCTAAGGTGTCTCTACGTGA 988
 DB 1161 TCACCATGACCAACGAGCTCTCGGCTCTCGAGCATCTCTGCCAAGGTGTCTCTATGTA 1220
 QY 989 AGGCAATGACATCTGGATGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
 DB 1221 AAGCCATTCGATTTGGATGGCAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
 QY 1049 ATGCTGCCATAAATTTTGTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAAGGC 1108
 DB 1281 ATGCTGCCCTTAACCTTTGTCTCGGCAACATAGGAGCTGCTCCGATTCAGAGGAAGC 1340
 QY 1109 AGAGCGCCCAAGCTTTGGAGGAAGATATCATCAAGAAAGTGTCTTCTATTTCCGCTGGCT 1168
 DB 1341 GGAGACATCAAGAGGAGTGAAGCT-----GGAGAAGCGCTTTAACTTCTCTGCT 1394
 QY 1169 ATGGCTTGGGCC---ACTGCTGAGGCAAGAGATGGAGGTCCATGGAAGGTTC---TG 1222
 DB 1395 ATGGATGGGGCCAGCCTGTCTACAGGCCAAGGATGGCATCTCAGTCAAGGGCGCCAACA 1454
 QY 1223 GCATTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGAAGAGAAACACCGCGAAAC 1282
 DB 1455 ACAGTAACACCAACACCCCTCTCTGACCATCTTAAGTCTCCAGAGGAGATGCGAAAC 1514
 QY 1283 TCTAGCTGACCTGAGCAAGAGAAATTGACCATCTCCCGGGGTGTCTTCCCTTTCACTT 1342
 DB 1515 TCTTCATCAGAGGGCCAGAGATCGAATAATATCCCGCATTTGGCTTCCCATGGCCT 1574
 QY 1343 TCTCATCTTCAATATCTTCTACTGGGTGTCTATAAAGTGTATGGTCAGAGATATCC 1402
 DB 1575 TCTCATTTTCAACATGTTCTACTGGATCATCTACAAGATTGTCCTGAGAGAGAGCTCC 1634
 QY 1403 ACCA 1406
 DB 1635 ACAA 1638

DEFINITION H. sapiens alpha-1 strychnine binding subunit of inhibitory glycine receptor mRNA.
ACCESSION X52009
VERSION X52009.1 GI:31850
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1715)
AUTHORS Grenningloh, G., Schmieden, V., Schofield, P.R., Seeburg, P.H., Siddique, T., Mohandas, T.K., Becker, C.M. and Betz, H.
TITLE Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes
JOURNAL EMBO J. 9 (3), 771-776 (1990)
MEDLINE 90183975
PUBMED 2155780
FEATURES
source
1..1715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xrefs="taxon:9606"
/chromosome="X"
/map="Xp21.2-p22.1"
/tissue_type="brain"
/dev_stage="fetal"
1..1715
/genes="GLYRA2"
297..1646
/genes="GLYRA2"
/notes="strychnine binding alpha-1 subunit"
/codon_start=1
/product="inhibitory glycine receptor"
/protein_id="CAA36258.1"
/db_xref="GI:31851"
/db_xref="GOA:P23415"
/db_xref="UniProt/Swiss-Prot:P23415"
/translation="MYSELTRLYLGSALVFSLAASKEAARSATKPMSPDFLDK
LMGRTSGYDARFPNFKGPPVNVSCNIFNSGSAETITMDYRVNIFLQWNRDLA
YNEYPDSDLOPMSLDLSWKPDLPFANEKGFHEITTDNKLRLSRNGVLSIRI
TUTLACPMDLNPFMDPTCIWQLESFGYTMNDLIPFEWQEGAVQVADGLTPQFLK
EKDLRYCTKHYNTFTCTIEARPHLERQMYLIQMYIPSLILVLSLWISFINMDA
APARVGLGITVLTMTQSSGRSASLPKSYVKAIDIWMAVCLLFFVSLALVAVNF
VSRQKELLRFRKRHRKDEBAGEGRNFSAYGMGPACLOAKDGLSVKGNNSNTN
PPAPSKSPEEMKLFQRAKKIDKISRIGFPMAFLIPNFWYLIYKIVREDVHQ"
ORIGIN
Query Match 40.0%; Score 656; DB 9; Length 1715;
Best Local Similarity 71.9%; Pred. No. 5e-169;
Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
QY 89 AGGAAGTCAATCTGNAACCAAGGGTCCAGCCCATGTCCTCTGATTCCTAGACA 148
DB 367 AGGAGGCTGAAGCTGCTCGCTCCGCAACCAAGCCTATGTACACCTCGGATTTCTGGATA 426
QY 149 AACTTATGGGCGCAACATCTCGGATATGATGCCAGGATTCGGCCCAATTTAAAGGCCAC 208
DB 427 AGCTAATGGGGAGAACCTCCGGATATGATGCCAGGATTCAGGCCCAATTTAAAGGTCCTC 486
QY 209 CCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAGCAACACCA 268
DB 487 CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGGTTCCATTTGCTGAGACCA 546
QY 269 TGACATACCGGGTGAATGTTCTTCTGGGCAACAGTGGAAATGACCAAGCTGCTCTACC 328
DB 547 TGGACTATAGGGTCAACATCTTCTCGCGGAGCAATGGAACGACCCCGCTGGCTATA 606
QY 329 GAGAATATCTGATGACTCTCTGACCTCGATCCCTCCATGCTGGACTCTATCTGGAAGC 388
DB 607 ATGAATACCTGACGACTCTCTGACCTGGACCCATCCATGCTGGACTCTCATCTGGAAC 666
QY 389 CAGACCTCTCTTTTGTCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCAAGCAACA 448

667 CTGACCTGTTCTTTCGCAACGAGAAAGGGGCCACTTCCATGAGATCACCAAGCAACA 726
QY 449 AGTTACTCGCATCTTCAAGATGGAATGTCTGACAGCATAGGCTGACCTCATTT 508
DB 727 AATTGCTAAGGATCTCCCGGAATGGAAATGCTCTTACAGCATAGAAATCACCCGTGAC 786
QY 509 TGTCTGCTGATGACCTCAAGAACTTCCCATGAGATCCAGACGTGCAAGATGACG 568
DB 787 TGGCTTGGCCCATGACCTTGAAGAAATTTCCCATGATGTCAGACATGATATCATGCAAC 846
QY 569 TTGAGAGCTCATCCATATCTCTGACGCCCTCTGCCATCTCTGTCATCTTCAGTTGGCTACA 628
DB 847 TGGAAAGCT-----TTGGATAATA 864
QY 629 CCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTCTGCTGCCAAGTGGCTGAGG 688
DB 865 CGATGAATGACCTCATCTTTGAGTGGAGGAACAGG---GAGCCGTGAGGTAGCAGATG 921
QY 689 GGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTCTTTTACCAAGC 748
DB 922 GACTAATCTGCCCCAGTTTATCTTGAAGGAAGAGAGGACTTGAGATACTGCACCAAGC 981
QY 749 ACTACAACACAGGGAATTCACCTGATCGAGGTAAAGTTTCACTTGGAAACGGCAGATGG 808
DB 982 ACTACAACACAGGTAAATTCACCTGATTTGAGGCCCGGTTCACCTGGAGCGCAGATGG 1041
QY 809 GCTACTATCTGATTCAGATGTACATCCACAGCTACTCATCTGCTCATCTCTGCTGCT 868
DB 1042 GTTACTACCTGATTCAGATGTATATTCACAGCTCTCATTTGTCATCTCTCATGATCT 1101
QY 869 CTTCTGGATCAACATGATGCTGCTGCCCTGCTGGGCTTGGGCTGAGCATCACCAAGCTGC 928
DB 1102 CTTCTGGATCAACATGATGCTGACCTGCTGCTGCTGGGCTTGGGCTAGGATCACCAAGCTGC 1161
QY 929 TCACATGACCAACAGAGCTCTGGCTCCCGGCTCTTTTGGCTTAAGGTGCTCTAGCTGA 988
DB 1162 TCACATGACCAACAGAGCTCTGGCTCTCGAGCATCTTCGCCCAAGGTGCTCTATGTA 1221
QY 989 AGGCAATCGACATCTGGATGGCTGTGCTCTCTTTGTTGCTGCTGCTGCTGCTGCTGCT 1048
DB 1222 AAGCAATTCACATTTGGATGGCAGTTTGGCTCTCTTTGTTGCTCAGCCCTATTAGAAT 1281
QY 1049 ATGCTGCCATAAATTTTGTCTCTGTCAGCAATAAAGAAATTCATACGACTTCGAAGAGGC 1108
DB 1282 ATGCTGCCTTAACTTTGTGCTCGGCAACATAAGAGCTGCTCCGATTCAGAGGAAGC 1341
QY 1109 AAGGGCCCAACGCTTTGGAGGAAGATATCATCCAGAAAGTCTGTTCTATTTCCGTGGCT 1168
DB 1342 GGAGACATCACAAAGGAGGATGAAGCT-----GGAGAAGGCCGCTTTAACTTCTCTGCT 1395
QY 1169 ATGGCTTGGGCC---ACTGCTCGAGGCAAGATGGAGGTCCATGGAAGGTTG---TG 1222
DB 1396 ATGGATGGGCCAGGCTGCTTACAGGCCAAGATGGCATCTCAGTCAAGGGGGCCCAACA 1455
QY 1223 GCATTTATAGTCCCAACCTCCAGCCCTCTTTTAAAGGAAGAGAAACACAGCGGAAC 1282
DB 1456 ACAGTAACACCAACACCCCTCTCTGACCATCTAAGTCCCCAGAGGAGATGCGAAGAAC 1515
QY 1283 TCTACGTGACCTGAGCCAAAGAGAAATTTGACCATCTCTCCGGGCTGCTCTCCCTTTCACTT 1342
DB 1516 TCTTATCATCAGAGGGCCAAAGAGATCGACAAATATCCCGCATTTGGCTTTCCCATGGCCT 1575
QY 1343 TCTCATCTTCAATCTCTTACTGGTGTCTATAAAGTGTATGTCAGAGATATCC 1402
DB 1576 TCTCATTTTCAACATGTTTCTACTGGATCATCTACAAATGTTGTCCTAGAGGACGCTCC 1635
QY 1403 ACCA 1406
DB 1636 ACNA 1639

RESULT 10

RNO310837	1359 bp	mRNA	linear	ROD 03-APR-2001
LOCUS				
DEFINITION	Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.			
ACCESSION	AJ310837			
VERSION	AJ310837.1	GI:13548660		
KEYWORDS	glycine receptor alpha 2 precursor.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
AUTHORS	Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivilotti, L.G.			
TITLE	Concentration dependence of single channel currents through rat recombinant alpha 1 glycine receptors			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1359)			
AUTHORS	Groot-Kormelink, P.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of Pharmacology, The School of Pharmacy, 29/39, Brunswick Square, London, WC1N 1AX, UNITED KINGDOM			
FEATURES	Location/Qualifiers			
source	1..1359			
	/organism="Rattus norvegicus"			
	/mol_type="mRNA"			
	/strain="Sprague-Dawley males"			
	/db_xref="taxon:10116"			
	/tissue_type="whole brain"			
	/clone_lib="rat brain marathon-ready cDNA"			
	/dev_stage="8-12 weeks"			
	1..1359			
	/function="ion channel"			
	/codon_start=1			
	/evidence="experimental"			
	/product="glycine receptor alpha 2 precursor"			
	/protein_id="CAC35981.1"			
	/db_xref="GOA:P22771"			
	/db_xref="GI:13548661"			
	/db_xref="UniProt/Swiss-Prot:P22771"			
	/translation="MNQLVNIITLAFPLGNTNHPAFCKDHSRSGKPSQTLSP SFLDKLGRSTGYDARIIPNPKSPVNVTCNIFINSFGSVETMDYVNIFLRQW NLSRLAYSYPDDLSDLDSPMLDSIWKDLPFANEXGANFHDVTDNKLRLSKNGKV LYSIRLTLSCPMDLNPFMDVQTCMQLESFGYTMNDLIFEWLSDGPVQVABGLTL POFLEKEBELGYCTKHYNTGKFTIEVRFHLERQMGYYLIQMYIPSLILVILSWVF WINMDAAPARVALGITVTLTMTTQSSSRASLPKVSIVKAIIDIMWAVCLLFFAALLE YAANFVSQHKFELRLRRKRONKEEDVTRERENFGYMGHCLQVKGDTAVKAT PANPLPQPKDADAIKKPFVDRAKRIDTISRAPPFLAFLIFNIFYWITYKIIRHEDVH KK"			
CDS				
	/function="ion channel"			
	/codon_start=1			
	/evidence="experimental"			
	/product="glycine receptor alpha 2 precursor"			
	/protein_id="CAC35981.1"			
	/db_xref="GOA:P22771"			
	/db_xref="GI:13548661"			
	/db_xref="UniProt/Swiss-Prot:P22771"			
	/translation="MNQLVNIITLAFPLGNTNHPAFCKDHSRSGKPSQTLSP SFLDKLGRSTGYDARIIPNPKSPVNVTCNIFINSFGSVETMDYVNIFLRQW NLSRLAYSYPDDLSDLDSPMLDSIWKDLPFANEXGANFHDVTDNKLRLSKNGKV LYSIRLTLSCPMDLNPFMDVQTCMQLESFGYTMNDLIFEWLSDGPVQVABGLTL POFLEKEBELGYCTKHYNTGKFTIEVRFHLERQMGYYLIQMYIPSLILVILSWVF WINMDAAPARVALGITVTLTMTTQSSSRASLPKVSIVKAIIDIMWAVCLLFFAALLE YAANFVSQHKFELRLRRKRONKEEDVTRERENFGYMGHCLQVKGDTAVKAT PANPLPQPKDADAIKKPFVDRAKRIDTISRAPPFLAFLIFNIFYWITYKIIRHEDVH KK"			
ORIGIN				
Query Match	39.9%; Score 654.8; DB 10; Length 1359;			
Best Local Similarity	70.8%; Pred. No. 1e-168;			
Matches	938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;			
QY	84	AAAAGAGGAGTCAAACTCTGGAAACCAAGGGGTCACAGCCATCTCCCTCTGATTTCTT	143	
DB	84	AGACCATGACTCCAGGCTCTGGAAAACATCCCTCGCAGACCCCTGCTCTCTCAGATTTCTT	143	
QY	144	AGCAAACTTATGGGCGACATCTGGATATGATGCGAGATTCGGCCCAATTTTAAAGG	203	
DB	144	GGATAAACTAATGGGAAGGACATCAGGATATGATGAAGAATCAGGCCAAATTTTAAAGG	203	
QY	204	CCCAACCGGTGAAGTCAACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCCACCAAGAC	263	
DB	204	TCCTCAGTAAAGCTTACTTGCAATATTTTATCAACAGTTTGGATCGGTACAGAAC	263	
QY	264	CACAATGGACTACCGGGTGAATCTCTTTTCGGGCAACAGTGAATGACCCAGCTGTG	323	
DB	264	CACCATGGACTACCGAGTGAACATTTTCTTGAGACAGCAGTGAACGATTCACGGCTGGC	323	
QY	324	CTACCCAGAAATATCTTGATGACTCTCTGGACCTCGATCCCTCCATCGTGGACTCTATCTG	383	
324	ATACAGTGAATACCCAGATGATTCCCTGGATTGGATCCATCGATCGTGGATTCGATTTG	383		
384	GAAGCAGACCTCTTTCTTTGCTAAATAGAAAGGGGCAACTTCCATGAGGTGACCAAGGA	443		
384	GAACCCAGATTTGTTCTTTGCTCAATAGAAAGGAGCAATTTCCATGATGTACCACTGA	443		
444	CAACAGTTACTGCGCATCTTCAAGAAATGGGAATGTGTGTACAGCATCAGGCTGACCTT	503		
444	TAAACAAGTTGCTGCGGATTTCCAAAATAGCAAAAGTGCTCTACAGTATTAGACTCACCTT	503		
504	CATTTGTCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAGACGTGCAACGAT	563		
504	GACTTATCTGCTCCATGACCTGAAGAACTTTCCAAATGGATGTCACACCTGTACAAAT	563		
564	GGAGCTTGAGAGCTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTTCAGTTGG	623		
564	GCAGCTGGAGATT-----TTGG	581		
624	CTACACCATGAAGACCTCGTCTTTGAGTGGCTGGAGATGCTCTCTGCTGTCCAACTGGC	683		
582	GTACACCATGAATGACCTGATATTTGAGTGGTTAAAGTGATGGTCT---CAGTACAAGTTGC	638		
684	TGAGGGGCTGACTCTGCCCCAGTTTATCTTGCGGGATGAGAAAGATCTAGGCTGTGTATC	743		
639	TGAGGACTCACCTCGCTCAGTTTATTTTGAAGAGAGAGAACTTGGCTATTGTCAC	698		
744	CAAGCACTACAAACACAGGGAAATTCACCTGCATCGAGGTAAAGTTTCACTGGAAACGCA	803		
699	AAAGCATTAACAACACTGGCAAGTTTACCTGCTGATTTAGAGTCAAGTTTCACTGGAAACGCA	758		
804	GATGGCTACTATCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG	863		
759	GATGGCTATTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	818		
864	GGTCTCTCTCTGATCAACATGATGCTGCCCTGCGCTGCGGCTGGGCTCGGCATCAACAC	923		
819	GGTCTCTCTCTGATCAACATGATGCTGCCCTGCGCTGCGGCTGGGCTCGGCATCAACAC	878		
924	CGTGTCTACCATGACCAACCCAGAGCTCTGGCTCCCGGGCTCTTTGCCCTAAAGTGTCTTA	983		
879	CGTCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	938		
984	CGTGAAGGCAATCGACATCTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1043		
939	TGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	998		
1044	GGAGTATGCTGCATATAATTTTGTCTCTGATGATGATGATGATGATGATGATGATGATG	1103		
999	GGATATGACGAGTGAATTTGTCTCCAGGCAACACAGGAGTCTCTGCTCTCGGAG	1058		
1104	AAGGCAGAGCGCCCAACCGCTTGGAGGAAGATATCATCCAAAGAAAGTCTGTTTCTATTTCG	1163		
1059	ACGACAGAAGAGGAGCAATTAAGGAAGATGTTACTCTGTAAGACCGTTTAACTTCAG	1118		
1164	TGGCTATGCTTGGGCACTGCTGCGGCAAGAGATGAGGTCCAAATGGAAGTTCG	1223		
1119	CGGTTATGGGATGGGTCACTGCTCCAGTGAAGATGATGATGATGATGATGATGATGATG	1175		
1224	CATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGGAAGGAGAAACCAACGGGAAACT	1283		
1176	-----ACCTGCAACCCACTTCGCAACCCCAAAAGATGCAAGATGCTATCAAGAGAA	1229		
1284	CTACGTGGAATGAGCCCAAGAGAAATGACACCACTCTCCCGGGCTGTCTTCCCTTTCACTTT	1343		
1230	GTTTGTGGATCGGCAAAAAGAAATCGACACCATATCTCGAGCTGCTTCCCACTGGCCTT	1289		
1344	CTCATCTTCAATATCTTCTACTGGGTGTCTTATAAAGTGTCTATGCTGAGAAATATCCA	1403		
1290	CCTCATTTTCAACATCTTTTACTGGATCACATCAAGATCATTCGGCATGAAGATGTCCA	1349		
1404	CCAG	1407		
1350	CAAG	1353		

```

RESULT 11
RNI6RAA2
LOCUS      RNI6RAA2               3160 bp    mRNA    linear    ROD 23-MAR-1993
DEFINITION R.norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.
ACCESSION  X61159
VERSION    X61159.1  GI:288344
KEYWORDS   inhibitory glycine receptor alpha subunit.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 3160)
AUTHORS   Kuhse,J., Kuryatov,A., Maulet,Y., Malosio,M.L., Schmieden,V. and
            Betz,H.
TITLE      Alternative splicing generates two isoforms of the alpha 2 subunit
            of the inhibitory glycine receptor
JOURNAL    FEBS Lett. 283 (1), 73-77 (1991)
MEDLINE    91243883
PUBMED     1645300
FEATURES   Location/Qualifiers
            source          1..3160
                        /organism="Rattus norvegicus"
                        /mol_type="mRNA"
                        /db_xref="taxon:10116"
                        532..1910
            CDS             /codon_start=1
                        /product="inhibitory glycine receptor alpha 2A subunit"
                        /protein_id="CAA43471.1"
                        /db_xref="GI:288345"
                        /db_xref="GOA:P22771"
                        /db_xref="UniProt/Swiss-Prot:P22771"
                        /translation="MNRQLNLTALFAFLGNHFRFAFKDHSRSKHPSQTLSP
                        SFLDKMRTSGYDRIEIRKPKFPPVNVTCNIFNSFGSVTETMDNKLRLSKNGKV
                        NDSRLAYSYPDSDLDPSMLDSIWKPDLPFANEKGNFHDVTTDNKLRLSKNGKV
                        LYSRLTSLSCPMDLKNFPMQVOTCTMOLESFGVTMDNLIPEWLSGDPVOVAEGLTL
                        POFILKEELKGVCTKHYNTGKFTCEVAFHLEKROMGYLIOMYIPSLILVILSVSEF
                        WINMDAPARVALGITVLTWTQSGSRASLPKYSYKAIIDWAWCLLFVPAALLE
                        YAAVNFVQKHFEFLRLRRQRONKEEDVTRSRPNFSGYGMGHCLQVKDGTAVAK
                        PANPLPQPPKADAIKKFVDRAKRIDTISRAAFPLAFILFNI FYWITYKIIRHEDVH
                        KK"
            sig_peptide     552..632
                        /note="inhibitory glycine receptor alpha 2A subunit"
            mat_peptide     631..1907
                        /product="inhibitory glycine receptor alpha 2A subunit"
ORIGIN
Query Match      39.9%; Score 654.8; DB 10; Length 3160;
Best Local Similarity 70.8%; Pred. No. 1.2e-168;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;

QY 84 AAAAGAGGAGTCAAACTCGAACCAGGGGTCCAGCCCATGTCCCTCTGATTTCT 143
DB 635 AGACCATGACTCCAGGTCTGGAACATCCCTCGCAGACCTGTCTCTTCAGATTCT 694
QY 144 AGCAAACTTATGGGCGACATCTCGATATGATGCCAGATTCGGCCCAATTTAAAG 203
DB 695 GGATAAATAATGGGAAGGACATCGAGATATGATGAAGAATCAGGCCAAATTTAAAG 754
QY 204 CCACCCCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAACAGAC 263
DB 755 TCCTCCAGTAAGCTTACTTGCATATTTTATCAACAGTTTGGATCGGTACAGAAAC 814
QY 264 CACAATGGACTACCGGGTGAATGCTCTTCTGGGCAACAGTGAATGACCCAGCTGTC 323
DB 815 CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGAAGATTCACGGCTGCG 874
QY 324 CTACCCAGATATCTCGATGACTCTCTGGACCTCGATCCCTCATCTGGACTCTATCTG 383
DB 875 ATACAGTGAATACCCAGATGATTCCTGGATTTGGATTCATCGATGCTGATTCGATTTG 934
384 GAAGCCAGACACTCTTTCTTTGCTAATAGAAAGGGGCCAACTTCCATGAGGTGACCAACGA 443
935 GAAACAGATTTGTTCTTTGCTCAATGAGAAAGAGCCAAATTTCCATGATGTACCACTGA 994
444 CAAACAGTTACTGCGCATCTTCAAGATGGGAATGTGCTGTACAGCATCAGGCTGACCCCT 503
995 TAAACAGTTGCTGCGGATTTCCAAAAATGGCAAGTGTCTTACAGTATTAGACTCACCTT 1054
504 CATTGTCCTGCTGCTGAGGACCTCAAGAACTTCCCATGGACATCCAGACCTGACAGAT 563
1055 GACTTTATCTGCTGCCATGGACCTCAAGAACTTTTCAATGGATGTCCAGACCTGTACAAT 1114
564 GCAGCTTGAGAGCTCATCTCATATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTCACTT 623
1115 GCAGCTGGAGATT-----TTGG 1132
624 CTACACCATGAAGACCTGCTGTTTGTAGTGGCTGGAAGATGCTCTCTGTGTTCAAGTGGC 683
1133 GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTGTATGGTC---CAGTACAAAGTTGC 1189
684 TGAGGGGCTGACTCTGCCCTGCTTTTATCTTGGGGATGAGAGGATCTAGGCTGTGTAC 743
1190 TGAAGGACTCACCTGCTGCTCAGTTTATTTTGAAGAGAGAGAACTTTGGCTATTGGCAC 1249
744 CAAGCACTACAAACAGAGGAAATTCACCTGCAATCGAGGTAAAGTTTTCACCTGGAACGCA 803
1250 AAGCATTTACAACACTGGCAAGTTTACCTGCAATGAGGTCAAGTTTCACTTGAAGCGCA 1309
804 GATGGGCTACTATCTGATTCAGATGTACATCCCAAGCCTTACTCATCGTGCATCCTGTCTGT 863
1310 GATGGGCTATTATTTGATTCAGATGTATATCCCAAGCCTGCTGATAGTCAATTTGTCTGT 1369
864 GGTCTCTCTTGGATCAACATGGATGCTGCCCTGCCGTGTGGGCTGGGCTGACCAAC 923
1370 GGTCTCTCTTGGATAAACATGGATGCTGCCAGGCTCTGCGAGGTTGCCCTTGGCATCACAAC 1429
924 CGTGTCTCACTGACCAACACCCAGAGCTCTGGCTCCCGGGCTCTTTGGCTTAAAGTGTCTTA 983
1430 CGTCTGCAATGACTACCCAGAGTTTCAAGTTCCAGGGCATCTCTGCCAAAGTCTCATTA 1489
984 CGTGAAGGCAATCGACATCTGGATGGCTGTGTGTCTCTTTGTTGTTGCTGCTGCTTGTGT 1043
1490 TGTGAAGCAATTGACATCTGGATGGCAGTGTGCTCTTCTGTTGTGTTGCTGCTTACT 1549
1044 CGAGTATCTGCTCAATAAATTTTGTCTGTCAGCATAAAGATTCATACGACTTTCGAG 1103
1550 GGAATATGCGAGTGAACCTTGTCTCCAGGCAACACAGAGGATTCCTTCGCTCCGGAG 1609
1104 AAGGCAGAGCGCCCAACCGCTTGGAGGAAGATATCATCAAGAAAGTCGTTTCTATTTCG 1163
1610 ACGACAGAAGAGCGCAGATAAGGAAGAGATGTTACTCGTGAAGCCGTTTAACTTCAG 1669
1164 TGGCTATGCTTGGGCGCATCTGCTGAGGCAAGATGAGAGTCCATAGGAAGTTCGCTG 1223
1670 CGGTTATGGATGGGTGCTGCTGCTCAAGTGAAGATGTTACAGCTGTCAAGGCTAC--- 1726
1224 CATTATAGTCTCCCAACCTCCAGCCCTTCTTCTAAGGGAAGGAGAAACCCACCGGAAC 1283
1727 -----ACCTGCAACCCACTTCCGCAACCCCAAAAGATGAGATGCTATCAAGAAGAA 1780
1284 CTACGTGGACTGAGCAAGAGAAATTTGACACCATCTCCCGGGCTGTCTTCCCTTTTCACTTT 1343
1781 GTTTGTGGATCGGCAAAAGAAATCGACACCATATCTCGAGCTGCTTCCCACTGSCCTT 1840
1344 CTTCAATCTTCAATATCTTCTACTGGGTGTCTATTAAGTGTCTATGTTGATGAGATATCCA 1403
1841 CCTCATTTTCAACATCTTTTACTGGATCACATACAGATCATTCGGCATGAAGATGTCCA 1900
1404 CCAG 1407
1901 CAAG 1904

```

```
RESULT 12
AX700251
LOCUS AX700251 3865 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 20 from Patent EP1284297.
ACCESSION AX700251
VERSION AX700251.1 GI:29536131
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1
AUTHORS Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1284297-A 20 19-FEB-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
source
1..3865
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/notes="Neonatal glycine receptor"
ORIGIN
Query Match 39.9%; Score 654.8; DB 6; Length 3865;
Best Local Similarity 70.8%; Pred. No. 1.2e-168;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
QY 84 AAAGAGGAGTCAAAATCTGGACCAAGGGTCCAGCCCATCTGCCCTCTGATTTCT 143
DB 1256 AGACCATGACTCCAGTCTGGAAAAATCCCTCGACAGCCTGTCTCTTCAGATTTCT 1315
QY 144 AGACAACTTATGGGCGAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAG 203
DB 1316 GGATAAATAATGGGAAGGACATCAGATATGATGCAAGATCAGGCCAAATTTTAAAG 1375
QY 204 CCCACCGTGAAGTCAACCTGGAACATCTTCATCAACAGTTTCAGTCCGTTCACCAAGAC 263
DB 1376 TCCCTCAGTAAACGTTACTTGCATATATTTTATCAACAGTTTGGATCGGTTCACAGAAC 1435
QY 264 CACAATGGACTACCGGTGAACTGCTTCTTGGGCAACAGTGAATGACCCAGCGCTGTC 323
DB 1436 CACCATGGACTACCGAGTGAACATTTCTTGAGACAGAGTGAAGATTCACGGCTGGC 1495
QY 324 CTACCGAGGATATCTGATGACTCTCTGGACCTCGATCCCTCATGCTGAGACTTATCTG 383
DB 1496 ATACAGTGAGTACCCAGATGATTCCTCGATTTGGATCCATCGATCGTGGATTCGATTG 1555
QY 384 GAAGCAGACCTCTCTTTGCTAATGAGAAAGGGGCAACTTCCATGAGTGACACCGA 443
DB 1556 GAAACCAAGATTTGCTTTTGGCAATGAGAAAGGAGGCAATTTCCATGATGTCCACACTGA 1615
QY 444 CAACAGTACTCGCATCTTCAAGATGGATGTGCTGACAGCATCAGGCTGACCT 503
DB 1616 TACAAGTGTCTCGGATTTCCAAAATGGCAAAGTCTCTACAGATTAGACTCACCTT 1675
QY 504 CATTTTGTCTGCTGATGACCTCAAGAACTTCCCATGGAATCCAGACGTGACAGAT 563
DB 1676 GACTTTATCTGTCCCATGACCTGAAGAACTTTCCAAATGGATGTCCAGACCTGTACAT 1735
QY 564 GCAGCTTGAGAGTCAATCCATACTCTGACGACCTCTGCGCATCTCTGTCACTTTCACTGG 623
DB 1736 GCAGCTGGAGAGTT-----TTGG 1753
QY 624 CTACACCATGAAGACTCGTGTGTTGAGTGGCTGGAAGATGCTCTGCTGTCCAGTGGC 683
DB 1754 GTACACCATGAATGACTGTATTTTTCAGTGGTGAAGTATGATGGTC---CAGTACAGTTGC 1810
QY 684 TGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTGTATC 743
DB 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAGAGAGAACTGGCTATTGGCAC 1870
```

```
744 CAAGCACTACACACAGGGGAATTCACCTGCATCGAGGTAAAGTTTCACCTCGAAGCGCA 803
1871 AAAGCATTTACACACTGGAAGTTTACCTGCAATGAGGTCAAGTTTCACTCGAAGCGCA 1930
804 GATGGCTACTATCTGATTCAGATGTACATCCCAAGCTTACTCATCGTCACTCTGCTCTG 863
1931 GATGGCTATTATTGATCCAGATGTATATCCCAAGCTTCTCATAGTCACTTTGCTCTG 1990
864 GGTCTCTTCTGGATCAACATGGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
1991 GGTCTCTTCTGGATCAACATGGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050
924 CGTGTCTCACCATGACCCAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
2051 CGTGTCTGACATGACTTACCCAGAGTTTCAAGTCTCCAGGCAATCTCTCCAAAGGCTCA 2110
984 CGTGAAGGCAATCGACATCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
2111 TGTGAAGCAATGACATCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170
1044 GGAGTATGCTGCCATAAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
2171 GGAATATGACAGTGAATTTGTCTCCAGGCAACACACAGGAGTTCCTTCTGCTCCGGAG 2230
1104 AAGGACAGGCGCCCAACGCTTGGAGGAAGATATCATCCAGAAAGTCTGTTCTATTCCG 1163
2231 ACGACAGAGAGGCGCAAGATTAAGGAAGAGATGTTACTCTGGAAGCGTTTAACTTCAG 2290
1164 TGGCTATGCTTGGGCGCACTGCTGAGGCAAGAGATGAGGTCCTCAATGGAAGTTCTGG 1223
2291 CGTTATGGATGGTCACTGCTCCAGTGAAGATGCTACAGTGTCAAGCTAC--- 2347
1224 CATTTATGTCCTCCAACTCTCCAGCCCTCTTTAAGGGAAGGAGAAACCAACGGGAACT 1283
2348 -----ACCTGCCAACCCACTTCCGCAACCCCAAGAGATGAGATGCTATCAAGAAGA 2401
1284 CTACGTGACTGACCAAGAGATTCACACCATCTCCGGGCTGCTTCCCTTTCACTTT 1343
2402 GTTGTGGATGGGCAAAAGAAATGACACCATATCTCGAGTGTCTTCCACTGGCCTT 2461
1344 CTTCTATTTCAATATCTTCTACTGGGTGCTATTAAGTGTCTATGATGCTGAGAAGATAT 1403
2462 CTTCTATTTCAATATCTTCTACTGGGTGCTATTAAGTGTCTATGATGCTGAGAAGAT 1403
1404 CCAG 1407
2522 CAAG 2525
```

```
RESULT 13
RNNEOGLY
LOCUS Rat NGlyR mRNA for neonatal glycine receptor.
DEFINITION Rat NGlyR mRNA for neonatal glycine receptor.
ACCESSION X57281
VERSION X57281.1 GI:56743
KEYWORDS glycine receptor; NGlyR gene.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 3865)
AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.
TITLE Cloning of a glycine receptor subtype expressed in rat brain and
spinal cord during a specific period of neuronal development
JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)
MEDLINE 91200276
PUBMED 1707830
REFERENCE 2 (bases 1 to 3865)
AUTHORS Hishinuma, F.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1991) F. Hishinuma, Mitsubishi Kasei Institute of
Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan
```

FEATURES
source

Location/Qualifiers
1. .3865
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="spinal cord"
/dev_stage="10 days old"
1. .3865
/gene="Ngyr"
1. .3865
/gene="Ngyr"
/evidence="NGlyR"
1173. .2531
/evidence="experimental"
/gene="NGlyR"
/codon_start=1
/product="neonatal type of glycine receptor"
/protein_id="CAA40549.1"
/db_xref="GI:56744"
/db_xref="GOA:P22771"
/db_xref="UniProt/Swiss-Prot:P22771"
/translation="NNRQIVNLTALFAFLGTHNREAFCKDHSRSGKHPQSLSL
SFLDKMGRTSGYDRIAPNFKGPPVNTCNIPINSFGSVTETMDIRVNFLEQOW
NNSRLATSLDPSMLDSIWKDPLFANEKGFHVDVTDNKLRLSKNGKV
LYSIRLTLSCFMDLKNFMDVQCTMQLESFGYTMNDLIIFEWLSDFVQVABGLT
POPILKEBELGCTKHYNTKFTCEVFKHLEKQMYLYLIQWYIPSLILVILSVSF
WINMDAAPARVALGITTTLTMTTQSSSRASLPKSYVKAIDIMWAVCLLFFVPAALLE
YAANVFSQKHEFLRLRRQKNEEDVTRERENFSYGMGHCLOVKDGTAVKAT
PANLPQPPADADAIKKFVDRANRDIITISRAFFLAFLIFNIFYITTYIKIRHEDVH
KK"

ORIGIN

Query Match 39.9%; Score 654.8; DB 10; Length 3865;
Best Local Similarity 70.8%; Pred. No. 1.2e-168;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
QY 84 AAAGAGGAGTCAAACTCTGGAAACCAAGGGGTCCAGCCCATGTCCTCCCTCTGATTTCTT 143
DB 1256 AGACCATGACTCCAGGTCTGGAAACATCCCTCGCAGACCTGCTCTCTTCAGATTTCTT 1315
QY 144 AGACAACTTATGGGGGAACATCTGGATATGATGCGAGGATTCGGGCCCAATTTTAAAGG 203
DB 1316 GGATAAATTAATGGGAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAGG 1375
QY 204 CCACCCGCTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTACCAAGAC 263
DB 1376 TCCTCAGTAAAGCTTACTTGGCAATATTTTATCAACAGTTTGGATCGGTACAGAAAC 1435
QY 264 CACAATGGACTACCGGTGAATGCTCTTTTGGCGCAACAGTGGAAATGACCCACGCTGTG 323
DB 1436 CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGGAAACGATTCACGGCTGGC 1495
QY 324 CTACCGAGATATCCTGATGACTCTCTGGACCTCGATCCCTCATCGTGGACTCTATCTG 383
DB 1496 ATACAGTGAGTACCGAGATGATTCCTGGATTTGGATCCATCGATGCTGGATTCGATTTG 1555
QY 384 GAAGCCAGACTCTCTCTTCTATGAGAAAGGGGCAACTTCCATGAGTGAACACCGGA 443
DB 1556 GAACACAGATTTGTTCTTTGCCAATGAGAAAGGACAAATTTCCATGATGTACCACTGA 1615
QY 444 CAACAAGTTACTCGGCATCTTCAAGAAATGGGAATGTCTGTACAGCATCAGGCTGACCCCT 503
DB 1616 TAACAAGTTCTCGGGATTTCCAAAATGCAAAAGTGTCTACAGATTAGACTCACTT 1675
QY 504 CATTTTGCTGCTGATGACCTCAAGAACTTCCCATGGACATCCAGACGTGACGAT 563
DB 1676 GACTTTATCTGCTCCATGACCTGAAGAACTTTTCAATGGATGTCCAGACCTGTACAAT 1735
QY 564 GCAGCTTGAGAGTCTATCCATCTCTGACGCCCTCTGCCATCTCTGTCTCAGTTGG 623
DB 1736 GCAGCTGGAGAGTT-----TTGG 1753
QY 624 CTACACCATGAAGACCTCGTTGTGAGTGGCTGGAAGATGCTCTCTGCTGCCAAGTGGC 683

DB 1754 GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTGTATGTC---CAGTACAAGTTCG 1810
QY 684 TCAGGGGCTGACTCTGCCCCAGTTTATCTTGGGATGAGAAGGATCTAGGCTGTGTAC 743
DB 1811 TGAAGGACTCACCCCTGCTCAGTTTATTTTGAAGAAGAAGAACTTGGCTATTGCAAC 1870
QY 744 CAAGCACTACAACACAGAGGAAATTCACCTGTCATCGAGGTAAAGTTTCACTCGAAGCGCA 803
DB 1871 AAGCATTAACAACATGGCAGTTTACCTGCAATGAGGTCAAGTTTCACTCGAAGCGCA 1930
QY 804 GATGGCTACTATCTGATTCAGATGATACATCCAGCCTTCTCATCGTGCATCCTGTCCTG 863
DB 1931 GATGGCTATTATTGATCCAGATGATATATCCAGCCTGCTGATAGTCAITTTTGTCTG 1990
QY 864 GCTCTCTCTCGATCAACATGATGCTGCCCTGCCCTGCTGGCTGGGCTGGGATCACCAAC 923
DB 1991 GGTCTCTCTTGGATAAACATGATGATGCTGCTGAGGTTGCCCTTGGGATCAACAAC 2050
QY 924 CGTGCTCACCATGACCAACAGAGCTCTGCTGCCGGGCTCTTTGCTTAAAGTGTCTTA 983
DB 2051 CGTCTGCAATGATACCCAGATTCAGTTCCAGGCGATCTCTGCCAAGGTCTCAT 2110
QY 984 CTTGAAGCAATCGACATCTGGATGGCTGTGTCTGCTCTTTTGTGCTGCTGCTGCTTGT 1043
DB 2111 TGTGAAGCAATTCACATCTGGATGGCAGTGTGCTCTTCTGTTTGTGCTGCTGCTTACT 2170
QY 1044 GGAGTATGCTGCCATAAATTTTCTGTCAGCATAAAGATTCATACAGACTTCGAAG 1103
DB 2171 GGAATATGAGCAGTGAACCTTTGCTCCAGGCAACACAAGGAGTCTCTGCTCCGGAG 2230
QY 1104 AAGGACAGGCGCCAAACGCTTGGAGAGATATCATCAAGAAAGTCGTTTCTATTTCG 1163
DB 2231 AGCACAAGAGGCGCAGATAAGAGAGATGTACTCTGTAAGCGGTTTAACTTCAG 2290
QY 1164 TGGCTATGCTTTGGGCTACTGCTGCGAGCAAGATGAGGTCCATGGAAGTTCGCTGG 1223
DB 2291 CGGTTATGGATGGGTCTACTGCTCCAGTGAAGATGTTACAGCTGTCAAGGCTAC--- 2347
QY 1224 CATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGAGAGGAGAAACCAACGCGAAGT 1283
DB 2348 -----ACCTGCCAACCCCTCCGCAACCCCAAAAGATGAGATGTATCAAGAAGAA 2401
QY 1284 CTACGTGGACTGAGCAAGAGAAATTCACACCATCTCCCGGGCTGTCTTCCCTTTTCACTTT 1343
DB 2402 GTTTGAGTGGCGCAAAAGATCGACACCATATCTCGAGCTGCTTCCCACTGSCCTT 2461
QY 1344 CTTCTATTTCAATATCTTCTACTGCTGCTTCTATTAAGTGTATGTCAGAGATATCCA 1403
DB 2462 CCTCATTTTCAACATCTTTTACTGGATCACATCAAGATCATTCGGCATGAAGATGCCA 2521
QY 1404 CCAG 1407
DB 2522 CAAG 2525
RESULT 14
RNO310834
LOCUS
DEFINITION
Rattus norvegicus mRNA for glycine receptor alpha 1 precursor,
primary transcript.
ACCESSION
AJ310834
VERSION
AJ310834.1 GI:13548654
KEYWORDS
alternative splicing; glycine receptor alpha 1 precursor.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus
REFERENCE
1
AUTHORS
Groot-Kormelink, P.J., Beato, P.J., Finotti, C., Harvey, R.J. and
Sivilotti, L.G.
TITLE
Achieving optimal expression for single channel recording: a

plasmid ratio approach to the expression of alpha1 glycine
receptors in HEK293 cells
J. Neurosci. Methods 113, 204-214 (2002)

Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.
Openings of the rat recombinant alpha 1 homomeric glycine receptor
as a function of the number of agonist molecules bound
J. Gen. Physiol. 119 (5), 443-466 (2002)

21977760
11981023

Burzomato, V., Groot-Kormelink, P.J., Sivillotti, L.G. and Beato, M.
Stoichiometry of recombinant heteromeric glycine receptors revealed
by a pore-lining region point mutation
Recept. Channels 9 (6), 353-361 (2003)

14698963

Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.
The activation mechanism of alpha1 homomeric glycine receptors
J. Neurosci. 24 (4), 895-906 (2004)

14749434

Groot-Kormelink, P.J.
Direct Submission
Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of
Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,
London, WC1N 1AX, UNITED KINGDOM
related splice variants AJ310835 and AJ310836.
Location/Qualifiers
1. .1350
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley males"
/db_xref="taxon:10116"
/tissue_type="whole brain"
/clone_lib="rat brain marathon-ready cDNA"
/dev_stage="8-12 weeks"
1. .1350
/function="ion channel"
/notes="alternative splicing, primary transcript"
/codon_start=1
/evidence=experimental
/product="glycine receptor alpha 1 precursor"
/protein_id="CAC35978.1"
/db_xref="GI:13548655"
/translation="MYSFNTLRFLYMETIVFVSLAASKEADAARSAPKPMSPDFLDK
LMRTSGYDARIRPNFPGPVNVCNIFINSGSIAETMDYRVNIFLRQWMDPRLA
VNEYDPDLSLDPSMLDSIKPDLPEANEKGFHEITTDNKLIRSRNGVLYSIRI
TITLACPMOLKHPMDVQICIMQLESFGTMDLIFEWQEGAVQVADGLTPQFLK
EKDRYCTKHYNTGKFTCEARFHLERQMYLIQMYIPSLIIVLSWISFWMNDA
APARVLGITTTLTMTQSSGRSLPKVSYVKAIDIMWAVCLLFVFSALLSYAVNF
VSRHKELLRRKRHHKDDGEGRFNFSAYMGPAQLQAKDGISVKGANNNTN
PAPAPSKSPFEMREKLFQRAKKDKISRIGFPMAPFLIFNMFYIIVREDVHNK"

Query Match 39.9%; Score 654.6; DB 10; Length 1350;
Best Local Similarity 72.6%; Pred. No. 1.2e-168;
Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;

QY 116 CCNAGCCCATGTCCTGATTTCTAGACAACTTATGGGGCAACATCTGGATATG 175
DB 98 CCAAGCCTATGTCACCTCGGACTTCTGGATAAGCTTATGGGAAGACTTCTGGGTATG 157

QY 176 ATGCCAGGATTCGGCCCAATTTTAAAGGCCACCCGTGAACGTGACCTGCAACATCTTCA 235
DB 158 ATGCCAGGATTCAGCCCAACTTTAAAGGTCTCTGTGAACGTGAGTTCGAACATCTTCA 217

QY 236 TCAACAGTTTCAGCTCCGTCACCAAGACCAATAGGACTACCGGGTGAATGTCTCTTGC 295
DB 218 TCAACAGCTTTGGTTCTATCGCGGAGACCAACCATGACCTACAGGGTCAACATCTTCTGTA 277

QY 296 GGCAACAGTGGAAATGATCCACCGCTGTCTACCGAGAATATCTGTGATGACTCTCTGGACC 355

278 GGCAGCAGTGGAAACGACCCCGCTCTCGCTCAATGAATATCTCTGACGACTCTCTGGACC 337
QY 356 TCGATCCCTCCATGTGGACTCTATCTGGAAGCCAGAACCTCTTTTGTGTAATAGAAAG 415
DB 338 TTGACCCATCCATGTGGATTCCATCTGGAAGCCTGACTTGTCTTTGCCAATGAGAAG 397
QY 416 GGGCCAACTTCCATGAGGTGACCGGACCAAGTTACTGCGCATCTTTCAAGAAATGGA 475
DB 398 GGGCCCACTTCCACGAGATCACCGGACCAAGCTGCTGAGAACTCTCCCGGAACGGCA 457
QY 476 ATGTGCTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGTGAGGACCTCAAGAACT 535
DB 458 AGTCTCTTACAGCATCAGAACTCAGCTGACTCTGGCTGCCCCATGGACCTCAAGAAAT 517
QY 536 TCCCATGGAATCAGACGTGACGATGACGATGAGAGCTGAGAGCTCATCCATCTCTGAGCC 595
DB 518 TCCCATGATGTACAGACATGTATCATGCAACTGGAAGCT----- 559
QY 596 CTCTGCCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAAGACCTCGTTTGTAGTGGC 655
DB 560 -----TTGGTTATACCATGAAACGACCTCATCTTTTGTAGTGGC 595
QY 656 TGGAAAGATCTCTCTGTCTCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGC 715
DB 596 ----AGAGCAAGAGCTGTGAGGTGGCAGAGGACTGACCTGCTCAGTTTATCTTGA 652
QY 716 GGGATGAGAAGGATCTAGGCTGTGTGACCAAGCCTTACAAACAGGAAATTCACCTGCA 775
DB 653 AGGAAGAGAAGGATCTGAGATATGCAACCAAGCCTTACAAACAGGTAATTCACCTGCA 712
QY 776 TCGAGGTAAGTTTACCTGGAAACGGCAGATGGGCTACTATCTGATTTCAGATGTATATCC 835
DB 713 TTGAGGCCGATTTCCACCTGGAAACGGCAGATGGGCTACTACCTGATCCAGATGTATATCC 772
QY 836 CCAGCTACTCTCATCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
DB 773 CCAGCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832
QY 896 CTGCGCGTGTGGGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
DB 833 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
QY 956 CCGGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
DB 893 CCGGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
QY 1016 GTCGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
DB 953 GCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
QY 1076 AGCATAAAGAAATTCATACGACTTTCGAAGAAGGAGGCGGCCAACGCTTTGGAGGAAGATA 1135
DB 1013 AACACAAGAACTC-----CTTCGATTTAGGAGGAAGCGGACATCAAGGATGATG 1066
QY 1136 TCAATCAAGAAAGTGTGTTTCTATTTTCCGTGGCTATGCTTGGGCTGCTGCTGCTGCTGCTGCT 1192
DB 1067 AGGGTGGAGAGGCGGCTTCAACTTCTCCGCTCTGAGGATGGGCGCCAGGCTGCTGCTCAAG 1126
QY 1193 CAAAGATGGAGGTCCAATGGAAGTTCGGAATTTATAGTCCCAACCTCCAGGCCCT- 1251
DB 1127 CCAAGATGGCATCTCTGTCAAGGGTGCCAAACAAACAAACACCAAGAACCCCGCTCTCTG 1186
QY 1252 --CTTCTAAGGGAAGGAGAAACCAACCGGAAACTCTACGTGGAGCTGAGCAAGAGAAATG 1309
DB 1187 CACCGTCCAAGTCCCGGAGGAGATCGGAAACTCTTTCATCCAGAGAGCAAGAGATCG 1246
QY 1310 ACACATCTCCCGGGTGTCTTCCCTTTTCACTTTTCTCTCATCTTCTCAATATCTTCTACTGG 1369
DB 1247 ACAAATATCTCGCATCGGTTTCCCATGCGCTTCTCTCATCTTCTCAATATCTTCTACTGGA 1306
QY 1370 TTGCTCTATAAAGTGTATGCTGAGAGATATCCACA 1406
DB 1307 TCATCTACAAGATCGTCCGAGAGAGGACGTCCACAA 1343

```

RESULT 15
LOCUS AY129229 1388 bp mRNA linear ROD 10-SEP-2003
DEFINITION Mus musculus strain C57BL/6 glycine receptor alpha 1 subunit
(Glra1) mRNA, complete cds; alternatively spliced.
ACCESSION AY129229
VERSION AY129229.1 GI:22535389
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1388)
AUTHORS Findlay,G.S., Phelan,R., Roberts,M.T., Homanics,G.E.,
Bergeson,S.E., Lopreato,G.F., Mihic,S.J., Blednov,Y.A. and
Harris,R.A.
TITLE Glycine receptor knock-in mice and hyperekplexia-like phenotypes:
comparisons with the null mutant
J. Neurosci. 23 (22), 8051-8059 (2003)
MEDLINE 22835873
PUBMED 12954867
REFERENCE 2 (bases 1 to 1388)
AUTHORS Lopreato,G.F., Findlay,G.S., Harris,R.A. and Mihic,S.J.
TITLE Mus musculus glycine receptor, alpha 1 subunit (Glra1)
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1388)
AUTHORS Lopreato,G.F., Findlay,G.S., Harris,R.A. and Mihic,S.J.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2002) Inst. of Cellular and Molecular Biology,
The Univ. of Texas at Austin, 2500 Speedway, Room 1.148, Austin, TX
78712, USA
FEATURES
source
1. .1388
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/tissue_type="whole brain"
1. .1388
/genes="Glra1"
13. .1362
/genes="Glra1"
/notes="alternatively spliced"
/codon_start=1
/product="glycine receptor alpha 1 subunit"
/protein_id="AAM97291.1"
/db_xref="GI:22535390"
translation="MYSFNTLRFYLWETIVFSLAASKEAARSAKPMSPSDFLDK
LMGRSGYDARIRPNPKGPPVNVSNIFINSFGSIAETMDYRVNIFLRQWNPRLA
YNEYPDSDLDPSMLDSLWKEDLPFANEKGAFHEITDNLKLRISRNGNVLISIRI
TITLACPMDLKNFPMVDQTCIWOLESFGYTMNDLIFEWQEGAVQVADGLTLPFLK
EKDLRYCKHYNTKFTCEIARFHLEQMYGLIQMYIPSLIIVILSWISFWINDA
APARVLYGTITVLMTTQSGSRASLPKVSIVKAIIDIMWAVCLLFVFSALLEYAAVNF
VSRHKELLERFRKRHHKDDSGEGEGRNFSVAGMPACLOAKDGI SVKGNNTTN
PPAPSKSFEEMRWKLFQRAKIDKISRIGFPMFLIFNMFWIYKIVRDEVHVK"
ORIGIN
Query Match 39.9%; Score 654.6; DB 10; Length 1388;
Best Local Similarity 72.6%; Pred. No. 1.2e-168;
Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;
QY 116 CCAGGCCATGTCCTCGATTCTCTAGACAACTTATGGGGCAACATCTGGATATG 175
DB 110 CCAAGCCTATGCACCCCTCGACTTCTCGATAAGCTCATGGGAGGACTTCTGGGTATG 169
QY 176 ATGCCAGGATTCCGCCCAATTTTAAAGCCACCCTGGAACGTGACCTGCAACATCTTCA 235
DB 170 ACGCCAGGATCAGCCCACTTTAAAGTCTCTCTGTGAATGTAAGTTGCAACATCTTCA 229
QY 236 TCACAGTTTCAGCTCCGTCACCAAGACCAATGCACTACCGGGTGAATGTTCTTCTGC 295
|||||

```

```

230 TCACAGTTTCGGTTCCTCGCTGAGACAAACCATGGACTATATAGGGTCAACATCTTCTCTGA 289
296 GGCAACAGTGGAAATGACCCACGCTGTCTCTACCGAGAATATCTCTGATGACTCTCTGGACC 355
290 GGCAAGTGGAAACACCCCGCTCTGGCTCAATGAATACCTGATGACTCTCTGGACC 349
356 TCGATCCCTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTGTGTAATGAGAAAG 415
350 TTGACCCATCTATGTTGGATTCCATCTGGAAGCCTGACTTGTTCCTTGGCAATGAGAAG 409
416 GGGCAACATTCATGAGGTGACACGGACAAACAAAGTTACTGCGCATCTTCAAGAATGGGA 475
410 GGGCCACTTCCACGAATACACAGCGAACAAACATGCTAAGAAATCTCCGGAATGGCA 469
476 ATGTGCTGTACAGCATCAGGCTGACCCCTCATTTTCTGCTGCTGATGGACCTCAAGAACT 535
470 ATGTCTCTTACAGCATCAGAATCACCTGACCTGGCTGGCCCATGGACCTGGAAGAAAT 529
536 TCCCATGGACATCCAGACGTGACAGTGCAGCTTTGAGAGCTCATCTCATCTCTGCAGCC 595
530 TCCCGATGGATGACAGAGCTGATCATGCAACTCGAAAGCT----- 571
596 CTCTGCCATCTCTGTCACTTTTCAGTTGGCTACCATGAAGAACCTCTGTTTGTAGTGGC 655
572 -----TTGGATATACCATGAACGACCTCATCTTTTGTAGTGGC 607
656 TGGAGAGTCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGC 715
608 AGGAGCAAG---GAGCTGTGCAGGTGGCAGCGGACTGACCTGCTCAGTTTATTTCTGA 664
716 GGGATGAGAAGGATCTAGGCTGTTGTACCAAGCACTACAAACAGGGAATTCACCTGCA 775
665 AGGAAGAGAAACCTGAGATACTGCACCAAGCACTACAAACAGGTAATTCACCTGCA 724
776 TCGAGTTAAAGTTTCACTCTGGAAACGGCAGATGGGCTATCTATCTGATTAGATGTATATCC 835
725 TCGAGGCCCGAATTCACCTGGAGCGGCGATGGGCTACTACCTGATCCAGATGTATATCC 784
836 CCAGCTACTCATCTGCTCTCTGCTGCTGCTGCTCTCTGATCAACATGATGCTGCTGCC 895
785 CCAGCTGCTCATCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844
896 CTGCCCCGTGTGGGCTCTGGGCTATCACCACCGTCTCACCATGATGATGATGATGATGAT 955
845 CAGCTCGTGTGGGCTGGGCTATCACCACAGTCTCACCATGATGATGATGATGATGATGAT 904
956 CCGGGCTCTTTGGCTTAAGTGTCTCTGTAAGAGCAATCGACATCTGGATGGCTGTGT 1015
905 CCGAGCCTCCCTTACCCCAAGGTGCTCTACGTGAAAGCTATTGACATTTGGATGGCTGT 964
1016 GTCTGCTCTTGTGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
965 GCTGCTCTTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
1076 AGCATAAAGAAATTCATACGACTTTCGAAGAGGAGGAGGCGCCCAAGCTTTGGAGGAAGA 1135
1025 AACACAAGGA-----ACTGCTTCGATTAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078
1136 TCATCAAGAAAGTGTGTTTCTATTTCCGTTGGCTATGCTGGCTGGGCTGCTGCTGCTGCT 1192
1079 AGGGTGGAGAGGCGGCTTCAACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1138
1193 CAAGAGATGGAGTCCAAATGGAAGGTTCTGGCAATTTATAGTCCC---CAACCTCCAGGCC 1249
1139 CCAAGATGGCATCTCTGTCAAGGGTGCACAAACAAACAAACAAACAAACAAACAAACAAAC 1198
1250 CTCTTTCTAAGGAAGAGAGAAACCAACCGGAAACTCTACGTGGAGCTGAGCCAAAGAAATTG 1309
1199 CGGCATCCAAAGTCCCGGAGGAGATGCGGAAACTCTTTCATCAGAGAGGAGGAGGAGGAG 1258
1310 ACACATCTCCCGGGGTGTCTTCCCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACT 1369
1259 ACAAGATATCTCGCATCGGTTTCCCATGGGCTTCTCTCATCTTCTCATCTTCTCAACAT 1318
|||||

```

Qy 1370 TTGCTATAAGTCCTATGTCAGAGATATCCACCA 1406
Db 1319 TCATCTATAAGATCGTCCGAGAGAGGATGTCCCAA 1355

Search completed: April 28, 2005, 14:18:14
Job time : 7725.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 01:49:33 ; Search time 1076.49 Seconds
(without alignments)
9018.523 Million cell updates/sec

Title: US-10-075-846-3
Perfect score: 1640
Sequence: 1 atgacaactctgttcctgc.....taaaaaaaaaaaaaaaaaaaaa 1640

Scoring table:
IDENTITY_NUC
Gapex 10.0 , Gapex 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s:*	2: Geneseqn1990s:*	3: Geneseqn2000s:*	4: Geneseqn2001s:*	5: Geneseqn2001bs:*	6: Geneseqn2002as:*	7: Geneseqn2002bs:*	8: Geneseqn2003as:*	9: Geneseqn2003bs:*	10: Geneseqn2003cs:*	11: Geneseqn2003ds:*	12: Geneseqn2004as:*	13: Geneseqn2004bs:*
--------------------	--------------------	--------------------	--------------------	---------------------	---------------------	---------------------	---------------------	---------------------	----------------------	----------------------	----------------------	----------------------

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1640	100.0	1640	6	AAL49660 Human HGR
2	1524.8	93.0	2585	6	AAL49659 Human gly
3	1521.2	92.8	1874	6	AAD33667 Human TRI
4	1503.8	91.7	1600	3	AAC61678 cDNA sequ
5	1004.6	61.3	1574	12	ADJ27240 Human TRI
6	947.8	57.8	1224	12	ADJ27254 Human TRI
7	817.8	49.9	1176	8	ACC46177 Human dit
8	807.8	49.3	993	6	AAL49665 Human gly
9	660.8	40.3	1746	10	ABV76108 Neuron-Sp
10	656	40.0	1715	13	ADQ89057 Human uro
11	654.8	39.9	3865	10	ADB85139 Rat neona
12	650.2	39.6	1857	8	ABZ18534 Group III
13	648.6	39.5	1386	10	ABA00851 GLRA rela
14	648.6	39.5	1854	10	ABA00850 GLRA rela
15	552.8	33.7	3069	10	ADC77686 Human 238
16	552.8	33.7	3069	10	ABA00854 GLRA rela
17	552.8	33.7	3069	13	ADQ89103 Human uro
18	551.8	33.6	2225	12	ADJ27249 Human TRI
19	521.4	31.8	2569	10	ABA00852 GLRA rela
20	520.8	31.8	1438	10	ABA00855 GLRA rela

21	479	29.2	679	6	AAL49661	Aal49661 Human gly
22	366.2	22.3	657	2	AAQ25722	AaQ25722 GABA-A re
23	299	18.2	554	4	AAD17170	Aad17170 Human ion
24	299	18.2	554	8	ACD01555	AcD01555 cDNA clon
25	299	18.2	554	10	ADB29246	AdB29246 Novel hum
26	288.8	17.6	1129	10	ABA00853	AbA00853 GLRA rela
27	253	15.4	253	10	ACA55673	AcA55673 Human sig
28	253	15.4	253	12	ADI55469	AdI55469 Human pol
29	232.6	14.2	2106	13	ADR24668	AdR24668 Breast ca
30	227.8	13.9	39796	3	AAC61681	Aac61681 Nucleotid
31	221	13.5	543	12	ACH76139	Ach76139 Human gen
32	221	13.5	2404	6	ABI99254	AbI99254 Mouse isc
33	219	13.4	492	9	ACH28631	Ach28631 Human adu
34	208	12.7	214	12	ACH89839	Ach89839 Human gen
35	204.4	12.5	1614	6	AAD22072	Aad22072 Dermacent
36	204.4	12.5	1614	6	AAD22070	Aad22070 Dermacent
37	202.8	12.4	1614	6	AAD22071	Aad22071 Dermacent
38	193.8	11.8	1628	10	ADP81662	AdP81662 Leukaemia
39	193	11.8	1555	2	AAT28989	Aat28989 Human GAB
40	191.4	11.7	1359	10	ADB78711	Adb78711 Human GAB
41	191.4	11.7	1359	10	ADB78693	Adb78693 Human GAB
42	191.4	11.7	1359	10	ADB78697	Adb78697 Human GAB
43	189.8	11.6	1359	6	ABK27334	Abk27334 Mutant ga
44	189.8	11.6	1359	10	ADB78712	Adb78712 Human GAB
45	189.8	11.6	1359	10	ADB78721	Adb78721 Human GAB

ALIGNMENTS

RESULT 1
AAL49660
ID AAL49660 standard; cDNA; 1640 BP.
XX
AC AAL49660;
DT 27-NOV-2002 (first entry)
XX
DE Human HGRA4 splice variant coding sequence SEQ ID NO: 3.
XX
KW Human; glycine receptor alpha subunit 4; HGRA4; HGRA4sv; splice variant;
KW cardiovascular disorder; reproductive disorder; neural disorder;
KW cardiac; antiarrhythmic; antianginal; antiidiarrheic; antiulcer;
KW neurotropic; neuroprotective; antibacterial; virucide; protozoacide;
KW nervous system disorder; gastrointestinal disorder; gene therapy;
KW infection; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200266606-A2.
XX
PD 29-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-US0004329.
XX
PR 16-FEB-2001; 2001US-0269535P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemers N;
PI Chang H;
XX
DR WPI; 2002-674925/72.
XX
DR P-PSDB; AA019187.
XX
PT New isolated nucleic acid molecules encoding human glycine receptor A4
PT (HGRA4) polypeptides, useful for preventing, treating and ameliorating
PT conditions, e.g. neural or gastrointestinal disorders.
XX
PS Claim 1; Fig 2; 349pp; English.
XX
CC The present invention provides the protein and coding sequences of the
human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.

29-AUG-2002.

13-FEB-2002; 2002WO-US004329.

16-FEB-2001; 2001US-0269535P.

(BRIM) BRISTOL-MYERS SQUIBB CO.

Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemers N;
Chang H;

WPI; 2002-674925/72.
P-PSDB; AAO19186.

New isolated nucleic acid molecules encoding human glycine receptor A4
(HGRA4) polypeptides, useful for preventing, treating and ameliorating
conditions, e.g. neural or gastrointestinal disorders.

Claim 14; Fig 1; 349pp; English.

The present invention provides the protein and coding sequences of the
human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.
The sequences can be used in the treatment of neural disorders,
gastrointestinal disorders, disorders related to hyperglycine receptor
activity, cardiovascular disorders, reproductive disorders, or bacterial,
viral and parasitic infections. The present sequence is a coding sequence
of the invention

Sequence 2565 BP; 596 A; 681 C; 590 G; 698 T; 0 U; 0 Other;

Query Match 93.0%; Score 1524.8; DB 6; Length 2565;
Best Local Similarity 97.3%; Pred. NO. 0;
Matches 1578; Conservative 0; Mismatches 2; Indels 42; Gaps 1;

Qy 1 ATGCAACTCTGTTCCTGGAAACCTCTCCCTCCTTCTCTCGAACCTGCCAGGCAG 60
Db 1 ATGCAACTCTGTTCCTGGAAACCTCTCCCTCCTTCTCTCGAACCTGCCAGGCAG 60

Qy 61 GTCTCTCCTCAGGGTGCCCTTGGCAAAGAGAAGTCAATCTGGAAACAAGGGGTCCCAG 120
Db 61 GTCTCTCCTCAGGGTGCCCTTGGCAAAGAGAAGTCAATCTGGAAACAAGGGGTCCCAG 120

Qy 121 CCCATGTCCCCCTCTGATTTCCTAGACAAAATTATGGGGCGAACAATCTGGATATGATGCC 180
Db 121 CCCATGTCCCCCTCTGATTTCCTAGACAAAATTATGGGGCGAACAATCTGGATATGATGCC 180

Qy 181 AGGATTGGCCCCCAATTTAAAGGCCACCCTGTAACGTGACCTGCAACATCTTCATCAAC 240
Db 181 AGGATTGGCCCCCAATTTAAAGGCCACCCTGTAACGTGACCTGCAACATCTTCATCAAC 240

Qy 241 AGTTTCAGCTCCGTCCACCAAGCACCAATGGAATACCGGGTGAATGTCTTTGGGGCAA 300
Db 241 AGTTTCAGCTCCGTCCACCAAGCACCAATGGAATACCGGGTGAATGTCTTTGGGGCAA 300

Qy 301 CAGTGGAAATGACCCACGCTGTCTTACCGAGAATACTCTGATGACTCTCTGACCTCGAT 360
Db 301 CAGTGGAAATGACCCACGCTGTCTTACCGAGAATACTCTGATGACTCTCTGACCTCGAT 360

Qy 361 CCTTCCATCTGGACTCTATCTGGAGCCAGACCTCTCTTTTGCTTAATGAGAAAGGGGCC 420
Db 361 CCTTCCATCTGGACTCTATCTGGAGCCAGACCTCTCTTTTGCTTAATGAGAAAGGGGCC 420

Qy 421 AAC TTCATGAGGTGACCCACGCAACAAGTTTACTGCGCATCTTCCAAGAAATGGGAATGTG 480
Db 421 AAC TTCATGAGGTGACCCACGCAACAAGTTTACTGCGCATCTTCCAAGAAATGGGAATGTG 480

Qy 481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGCTGATGGACCTCAAGAACTTCCCC 540
Db 481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGCTGATGGACCTCAAGAACTTCCCC 540

Qy 541 ATGGACATCCAGAGGTGACGATGAGCTCATCAATCTCTGACGCCCTCTG 600
Db 541 ATGGACATCCAGAGGTGACGATGAGCTCATCAATCTCTGAGACTTCTGAGACTT----- 577

Qy	601	CCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAGAAGACCTCTGTGTTGAGTGGCTGGAA	660
Db	578	-----TTGGCTACACCATGAAGAAGACCTCTGTGTTGAGTGGCTGGAA	618
Qy	661	GATGCTCTCTGTCTCAAGTGGCTGAGGGGCTGACCTCTGCCCCAGTTTATCTTTGCGGGAT	720
Db	619	GATGCTCTCTGTCTCAAGTGGCTGAGGGGCTGACCTCTGCCCCAGTTTATCTTTGCGGGAT	678
Qy	721	GAGAAGGATCTAGGCTGTGTACCAAGCACTACAACACAGGGAAATTCACCTGCATCGAG	780
Db	679	GAGNAGGATCTAGGCTGTGTACCAAGCACTACAACACAGGGAAATTCACCTGCATCGAG	738
Qy	781	GTAAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTACATCCCCAGC	840
Db	739	GTAAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTACATCCCCAGC	798
Qy	841	CTACTCATCTCATCTCTGCTCTGGGTCTCCTTCTGGATCAACATGGATGCTGCCCTGCC	900
Db	799	CTACTCATCTCATCTCTGCTCTGGGTCTCCTTCTGGATCAACATGGATGCTGCCCTGCC	858
Qy	901	CGTGTGGGCTGGGCATCACACCGTCTCACCATGACCAACCCAGAGCTCTGGCTCCCGG	960
Db	859	CGTGTGGGCTGGGCATCACACCGTCTCACCATGACCAACCCAGAGCTCTGGCTCCCGG	918
Qy	961	GCCTCTTTGGCTTAAGGTGTCTACGTGAAGGCAATCGACATCTGGATGGCTGTGTCTG	1020
Db	919	GCCTCTTTGGCTTAAGGTGTCTACGTGAAGGCAATCGACATCTGGATGGCTGTGTCTG	978
Qy	1021	CTCTTTGTGTTCCGTGCTTGGTGTATGCTGCCATAAATTTTGTCTTCGTACGAT	1080
Db	979	CTCTTTGTGTTCCGTGCTTGGTGTATGCTGCCATAAATTTTGTCTTCGTACGAT	1038
Qy	1081	AAGAATTCATACGACTTCGAAGAAGCAGAGCGCCACCGCTTGGAGGAAGATATCATC	1140
Db	1039	AAGAATTCATACGACTTCGAAGAAGCAGAGCGCCACCGCTTGGAGGAAGATATCATC	1098
Qy	1141	CAAGAAAGTCGTTTCTATTTTCCGTGGCTATGGCTTGGGCCACTCCCTGCAGGCAAGAT	1200
Db	1099	CAAGAAAGTCGTTTCTATTTTCCGTGGCTATGGCTTGGGCCACTCCCTGCAGGCAAGAT	1158
Qy	1201	GGAGGTCCAATGGAAGGTTCTGGCATTTATAGTCCCCAACTCCAGCCCCCTCTTCTAAGG	1260
Db	1159	GGAGGTCCAATGGAAGGTTCTGGCATTTATAGTCCCCAACTCCAGCCCCCTCTTCTAAGG	1218
Qy	1261	GAAGGAAACCAACGCGGAACTCTACGTGCACTGAGCCAAAGAGATTGACACCATCTCC	1320
Db	1219	GAAGGAAACCAACGCGGAACTCTACGTGCACTGAGCCAAAGAGATTGACACCATCTCC	1278
Qy	1321	CGGCTCTCTTCCCTTTTCACTTTTCCCTCATCTTCAATATCTTCTACTGGGTGTCTATAA	1380
Db	1279	CGGCTCTCTTCCCTTTTCACTTTTCCCTCATCTTCAATATCTTCTACTGGGTGTCTATAA	1338
Qy	1381	GTGCTATGGTCAGAAGATATCCACAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCCT	1440
Db	1339	GTGCTACGGTCAGAAGATATCCACAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCCT	1398
Qy	1441	GCTGCTGGCCTCTGCTTCTCTCTGGTGGGCTTTCTCCCTCAGTTAGACTCCATTAGGG	1500
Db	1399	GCTGCTGGCCTCTGCTTCTCTCTGGTGGGCTTTCTCCCTCAGTTAGACTCCATTAGGG	1458
Qy	1501	GTTTGGACAGTTTCTTCTCTGATCTCCCACTCAGAACTTCAACTACCAAGTCCCAAGCTAT	1560
Db	1459	GTTTGGACAGTTTCTTCTCTGATCTCCCACTCAGAACTTCAACTACCAAGTCCCAAGCTAT	1518
Qy	1561	GTGGGCCCTATATGTCATGGTGCCAATGGTGGCTGTACTTATAAGATGGCTTATCTACCC	1620
Db	1519	GTGGGCCCTATATGTCATGGTGCCAATGGTGGCTGTACTTATAAGATGGCTTATCTACCC	1578
Qy	1621	TA 1622	
Db	1579	TA 1580	

1621 TA 1622

1579 TA 1580

241	Db	AGTTTCAGCTCCATCA	CCAAAGACCA	CAATGGACTAC	CGGGTGAA	TGCTTTCTTTGGCGCA	300	
301	Qy	CAGTGGAA	TGACCCACAGCCCTGTCT	CTACCGAGAA	TATCCTGATGACTCT	CTGGACCTCGAT	360	
301	Db	CAGTGGAA	TGACCCACAGCCCTGTCT	CTACCGAGAA	TATCCTGATGACTCT	CTGGACCTCGAT	360	
361	Qy	CCCTCC	ATGCTGGACTCT	ATCTGGAAGCC	AGACCTCTTTCTGCT	TAATGAGAAGGGCC	420	
361	Db	CCCTCC	ATGCTGGACTCT	ATCTGGAAGCC	AGACCTCTTTCTGCT	TAATGAGAAGGGCC	420	
421	Qy	AAC	TTCCATGAGGTG	ACCA	CGGACAA	CAAGTTACTGCGCATCT	CAAGATGGGAATGTG	480
421	Db	AAC	TTCCATGAGGTG	ACCA	CGGACAA	CAAGTTACTGCGCATCT	CAAGATGGGAATGTG	480
481	Qy	CTGT	TACAGCATCAGGCTG	ACCTCATTTTGTCT	GTGCTGATGACCT	CAAGAACTTCCCC	540	
481	Db	CTGT	TACAGCATCAGGCTG	ACCTCATTTTGTCT	GTGCTGATGACCT	CAAGAACTTCCCC	540	
541	Qy	ATG	GACATCCAGACG	TGCAGTGAC	CTTGTGAGAGCT	CATCTCTGTCGACCCCTCTG	600	
541	Db	ATG	GACATCCAGACCT	CTCAGATG	ACGCTTGTGAGAGT	-----	577	
601	Qy	CCAT	CTCTGTCA	CHTTTCAGTTGGCT	CACCATGAA	AGACCTCGTGTGTTGAGTGGCTGGAA	660	
578	Db	-----	-----	-----	-----	-----	618	
661	Qy	GAT	GCTCCTGCTG	TCCAAAGTGGCTG	AGGGCTGACTCTG	CCCCAGTTTATCTTTGCGGGAT	720	
619	Db	GAT	GCTCCTGCTG	TCCAAAGTGGCTG	AGGGCTGACTCTG	CCCCAGTTTATCTTTGCGGGAT	678	
721	Qy	GAGA	AGGATCTTAGGCTGT	TGTACCAAGCACT	CAACACAGG	GAAATTCACCTGCATCGAG	780	
679	Db	GAGA	AGGATCTTAGGCTGT	TGTACCAAGCACT	CAACACAGG	GAAATTCACCTGCATCGAG	738	
781	Qy	GTA	AGTTTCACTG	GGAACGGCAGATGGG	CTACTATCTGAT	TCAGATGATACCCCGCAGC	840	
739	Db	GTA	AGTTTCACTG	GGAACGGCAGATGGG	CTACTATCTGAT	TCAGATGATACCCCGCAGC	798	
841	Qy	CTA	CTCATCGT	CATCTGCTCTG	GGTCTCTTCTG	GATCAACATGGATGCTGCCCTGCC	900	
799	Db	CTA	CTCATCGT	CATCTGCTCTG	GGTCTCTCTTCTG	GATCAACATGGATGCTGCCCTGCC	858	
901	Qy	CGT	GTGGGCTGGGCAT	CACACCGT	GCTCACCATGAC	CCACCCAGAGCTCTGGCTCCCGG	960	
859	Db	CGT	GTGGGCTGGGCAT	CACACCGT	GCTCACCATGAC	CCACCCAGAGCTCTGGCTCCCGG	918	
961	Qy	GCCT	CTTTGGCTAAGT	TGCTTACGTGAAG	CAATCGACATCTG	GATGGCTGTGTGTCTG	1020	
919	Db	GCCT	CTTTGGCTAAGT	TGCTTACGTGAAG	CAATCGACATCTG	GATGGCTGTGTGTCTG	978	
1021	Qy	CTCT	TTGTGTTGCTGCTG	CTGGATGCTG	CCATAATTTGTTTCT	CTGTCAGCAT	1080	
979	Db	CTCT	TTGTGTTGCTGCTG	CTGGATGCTG	CCATAATTTGTTTCT	CTGTCAGCAT	1038	
1081	Qy	AA	GAAATTCATAC	GACTTCGAAG	AGCAGAGGCGCCAA	CGCTTGGAGGAAGATATCATC	1140	
1039	Db	AA	GAAATTCATAC	GACTTCGAAG	AGCAGAGGCGCCAA	CGCTTGGAGGAAGATATCATC	1098	
1141	Qy	CA	GAAAGT	CGTTTCTATTTCCG	TGGCTATGCTTGGGCC	CACTGCTGCAAGGAAGAT	1200	
1099	Db	CA	GAAAGT	CGTTTCTATTTCCG	TGGCTATGCTTGGGCC	CACTGCTGCAAGGAAGAT	1158	
1201	Qy	GG	AGTCC	CAATGGAA	GGTTCTGGCA	TTTATAGTCCCAACCTCCAGG	CCCTCTTCTAAGG	1260
1159	Db	GG	AGTCC	CAATGGAA	GGTTCTGGCA	TTTATAGTCCCAACCTCCAGG	CCCTCTTCTAAGG	1218
1261	Qy	GA	AGGAGAA	CCACCGG	AAACTCTA	CGTGGACTGAGC	CAAGAGAATTGACACCATCTCC	1320
1219	Db	GA	AGGAGAA	CCACCGG	AAACTCTA	CGTGGACTGAGC	CAAGAGAATTGACACCATCTCC	1278
1321	Qy	C	GGGCTGTCT	CCCTTCACTTT	CTCATCTT	CAATATCTTCTA	CTGGGTGTGTATAA	1379
1379	Db	CA	GGGCGAA	TTCCCTTCACTTT	CTCATCTT	CAATATCTTCTA	CTGGGTGTGTATAA	1338

5 T.J. PEA

RESULT 5
ADJ27240

ADJ27240
ID ADJ27240 standard: DNA: 1574 BP

XX
AC

XX
DT 20-MAY-2004 (first entry)

XX DE Human TRICH-24 coding sequence, SEQ ID 72.

Human; Transporters and Ion Channels protein; TRICH; Transporter;
 Ion Channel; transport disorder; neurological disorder; muscle disorder;
 Immunological disorder; cell proliferative disorder; infection;
 Neuroprotective; Muscular; Respiratory; Antidiabetic; Antiparkinsonian;
 Nootropic; Anticonvulsant; Antithyroid; Ophthalmologic; Antiparasitic;
 Antiinflammatory; Cardiant; Antitanginal; Antiarrhythmic; Antiasthmatic;
 Endocrine; Hypotensive; Anti-HIV; Antianemic; Anticancerotic;
 Dermatological; Antiatheritic; Osteopathic; Antipsoriatic; Rheumatic;
 Immunosuppressive; Antitumor; Gastrointestinal; Hepatotropic; Cytostatic;
 Virucide; Fungicide; Antiparasitic; Protozoacide; Antihelminthic;
 Gene Therapy. TRICH-24: gene: ds.

XX Homo sapiens.

AA WO2004013293-A2. PN

12-FEB-2004.

01-AUG-2003; 2003WO-US024164.

AA 02-AUG-2002; 2002US-0400953P. PR

PK 12-AUG-2002; 2002US-040332IF
PR 19-AUG-2002; 2002US-0404693P.

PR 29-AUG-2002; 2002US-040706ZF:
XX

PA (INCY-) INCYTE CORP.
XX

PI Richardson TW, Hata
PI Lee CV Chawla NK

PI Becha SD, Mason PM, Jiang X, Jackson AA, Marquis JP, Das D;
PI Gietzen KJ, Ding L, Baughn MR, Lee EA, Thornton MB, Borowsky ML;
PI Harland L, Arvizu CS, Yao MG, Gandhi AR, Au-Young JK, Warren BA;
PI Lal PG, Lindquist EA;

XX
DR WPT: 2004-180430/17.

DR WPI; 2004-180430/
DR P-PSDB: ADJ27192.

xx New human transporters and ion channels and nucleic acids encoding them,
PT useful for diagnosing, treating or preventing cystic fibrosis, Becker's
PT muscular dystrophy, Parkinson's disease, myasthenia gravis, prion diseases
PT or cancers.

XX	Claim 5; SEQ ID NO 72; 357pp; English.
PS	The present invention relates to novel human Transporters and Ion
XX	Channels proteins referred to collectively as TRICH (TRICH-1 - TRICH-48,
CC	ADJ27169-ADJ27216) and their coding sequences (ADJ27217-ADJ27264). The
CC	TRICH proteins and coding sequences are useful in the diagnosis,
CC	treatment and prevention of transport, neurological, muscle,
CC	immunological or cell proliferative disorders, e.g. akinesia, amyotrophic
CC	lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's
CC	muscular dystrophy, Bell's palsy, diabetes mellitus, diabetes insipidus,
CC	diabetic neuropathy, Parkinson's disease, malignant hyperplasia,
CC	myasthenia gravis, peripheral neuropathy, Alzheimer's disease,
CC	Huntington's disease, Grave's disease, cataracts, bacterial and viral
CC	meningitis, brain abscess, prion disease, Creutzfeldt-Jakob disease,
CC	fatal familial insomnia, neurofibromatosis, Down syndrome,
CC	cardiomyopathy, myocarditis, myotonic dystrophy, lipid myopathy, angina,
CC	arrhythmias, asthma, Cushing's syndrome, hypertension, hypoglycaemia,
CC	epilepsy, lactic acidosis, AIDS, anaemia, adult respiratory distress
CC	syndrome, atherosclerosis, bronchitis, Crohn's disease, atopic
CC	dermatitis, erythroblastosis fetalis, osteoarthritis, osteoporosis,
CC	pneumocystis, psoriasis, rheumatoid arthritis, scleroderma, systemic
CC	lupus erythematosus, ulcerative colitis, cirrhosis, cancers, viral,
CC	bacterial, fungal, parasitic, protozoal or helminthic infections.
XX	
SQ	Sequence 1574 BP; 364 A; 418 C; 402 G; 390 T; 0 U; 0 Other;
	Query Match 61.3%; Score 1004.6; DB 12; Length 1574;
	Best Local Similarity 83.2%; Pred. No. 3.8e-277;
	Matches 1342; Conservative 0; Mismatches 57; Indels 262; Gaps 7;
QY	1 ATGACAACTCTTGTTCTCGCAACCCTCTCCTTCTTCTCTGGACCCCTGCCAGGGCAG 60
DB	219 ATGACAACTCTTGTTCTCGAAGCCTCTCTTCTTCTTCTGGACCCCTGCCAGGGCAG 278
QY	61 GTCTCTCTCAGGTGGCCCTTGGCAAAGAGGAAGTCATAATCTGGAACCAAGGGGTCCAG 120
DB	279 GTCTCTCTCAGGTGGCCCTTGGCAAAGAGGAAGTCATAATCTGGAACCAAGGGGTCCAG 338
QY	121 CCCATGTCCCCTCTCATTTCTTAGACAAACTATTGGGGCGGAACATCTGGATATGATGCC 180
DB	339 CCCATGTCCCCCTCTCATTTCTTAGACAAACTATTGGGGCGGAACATCTGGATATGATGCC 398
QY	181 AGGATTCGGGCCCAATTTTAAAGGCCCCACCCGTGAACGTGACCTGCACCAATCTTCATCAAC 240
DB	399 AGGATTCGGGCCCAATTTTAAAGGCCCCACCCGTGAACGTGACCTGCACCAATCTTCATCAAC 458
QY	241 AGTTTCAGTCCGTACCAAGAACCAATPGACATACCGGGTGAATGTTCTTTGGGGCAA 300
DB	459 AGTTTCAGTCCGTACCAAGAACCAATPGACATACCGGGTGAATGTTCTTTGGGGCAA 518
QY	301 CAGTGAANTGACCCAGCCTGTCTACCGAGAATATCTGTGATCTCTCGACCTCGAT 360
DB	519 CAGTGAANTGACCCAGCCTGTCTACCGAGAATATCTGTGATCTCTCGACCTCGAT 578
QY	361 CCCTCCATGCTGACCTCTATCTGGAAGCCAGACCTCTTCTTTGCTTAATGAGAAAGGGGCC 420
DB	579 CCCTCCATGCTGACCTCTATCTGGAAGCCAGACCTCTTCTTTGCTTAATGAGAAAGGGGCC 638
QY	421 AACTTCCATGAGGTGACCAAGCAACAAGTTACTGCGCATCTTCAAGAAATGGGAATGTG 480
DB	639 AACTTCCATGAGGTGACCAAGCAACAAGTTACTGCGCATCTTCAAGAAATGGGAATGTG 698
QY	481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCC 540
DB	699 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCC 758
QY	541 ATGGACATCCAGACGTGACGATGACGTTGAGAGCTCATCTACTCTGACGCCCTCTG 600
DB	759 ATGGACATCCAGACGTGACGATGACGTTGAGAGCT----- 795
QY	601 CCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAAGACCTCGTGTGTTGAGTGGCTGAA 660

RESULT 6
ADJ27254
ID ADJ
XX
AC ADJ
XX

DT	20-MAY-2004	(first entry)	Best Local Similarity 86.1%; Pred. No. 6.7e-261;		
XX	Human	TRICH-38 coding sequence, SEQ ID 86.	Matches 1142; Conservative 0; Mismatches 2; Indels 183; Gaps 1;		
XX	Human	TRICH-38 coding sequence, SEQ ID 86.			
KW	Human; Transporters and Ion Channels protein; TRICH; Transporter;				
KW	Ion Channel; transport disorder; neurological disorder; muscle disorder;				
KW	Immunological disorder; cell proliferative disorder; infection;				
KW	Neuroprotective; Muscular; Respiratory; Antidiabetic; Antiparkinsonian;				
KW	Nootropic; Anticonvulsant; Antihypertensive; Ophthalmological; Antibacterial;				
KW	Antiinflammatory; Cardiac; Antianemic; Antiarthritic; Antidiabetic;				
KW	Endocrine; Hypotensive; Anti-HIV; Antianemic; Antiarthritic; Antidiabetic;				
KW	Dermatological; Antiarthritic; Osteoporotic; Antipsoriatic; Antirheumatic;				
KW	Immunosuppressive; Antitumor; Gastrointestinal; Hepatotropic; Cytostatic;				
KW	Virucide; Fungicide; Antiparasitic; Protozoacide; Antihelminthic;				
KW	Gene Therapy; TRICH-38; Gene; ds.				
OS	Homo sapiens.				
XX					
PN	WO2004013293-A2.				
PD	12-FEB-2004.				
XX					
PF	01-AUG-2003; 2003WO-US024164.				
PR	02-AUG-2002; 2002US-0400953P.				
PR	12-AUG-2002; 2002US-0403321P.				
PR	19-AUG-2002; 2002US-0404693P.				
PR	29-AUG-2002; 2002US-0407062P.				
XX	(INCY-) INCYTE CORP.				
PA					
XX	Richardson TW, Hafalia AJA, Khare R, Tran UK, Emerling BM;				
PI	Lee SY, Chawla NK, Elliott VS, Swarnakar A, Chang H, Kable AE;				
PI	Becha SD, Mason PM, Jiang X, Jackson AA, Marquis JP, Das D;				
PI	Gietzen KJ, Ding L, Baughn MR, Lee EA, Thornton MB, Borowski ML;				
PI	Harland L, Arvizu CS, Yao MG, Gandhi AR, Au-Young JK, Warren BA;				
PI	Lal PG, Lindquist EA;				
XX	-WPI: 2004-180430/17.				
DR	P-PSDB; ADJ27206.				
XX					
PT	New human transporters and ion channels and nucleic acids encoding them,				
PT	useful for diagnosing, treating or preventing cystic fibrosis, Becker's				
PT	muscular dystrophy, Parkinson's disease, myasthenia gravis, prion disease				
PT	or cancers.				
XX					
PS	Claim 5; SEQ ID NO 86; 357pp; English.				
XX					
CC	The present invention relates to novel human Transporters and Ion				
CC	Channels proteins referred to collectively as TRICH (TRICH-1 - TRICH-48,				
CC	ADJ27169-ADJ27216) and their coding sequences (ADJ27217-ADJ27264). The				
CC	TRICH proteins and coding sequences are useful in the diagnosis,				
CC	treatment and prevention of transport, neurological, muscle,				
CC	immunological or cell proliferative disorders, e.g. akinesia, amyotrophic				
CC	lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's				
CC	muscular dystrophy, Bell's palsy, diabetes mellitus, diabetes insipidus,				
CC	diabetic neuropathy, Parkinson's disease, malignant hyperplasia,				
CC	myasthenia gravis, peripheral neuropathy, Alzheimer's disease,				
CC	Huntington's disease, Grave's disease, cataracts, bacterial and viral				
CC	meningitis, brain abscess, prion disease, Creutzfeldt-Jakob disease,				
CC	fatal familial insomnia, neurofibromatosis, Down syndrome,				
CC	cardiomyopathy, myocardiitis, myotonic dystrophy, lipid myopathy, angina,				
CC	arrhythmias, asthma, Cushing's syndrome, hypertension, hypoglycaemia,				
CC	epilepsy, lactic acidosis, AIDS, anaemia, adult respiratory distress				
CC	syndrome, atherosclerosis, bronchitis, Crohn's disease, atopic				
CC	dermatitis, erythroblastosis fetalis, osteoarthritis, osteoporosis,				
CC	pancreatitis, psoriasis, rheumatoid arthritis, scleroderma, systemic				
CC	lupus erythematosus, ulcerative colitis, cirrhosis, cancers, viral,				
CC	bacterial, fungal, parasitic, protozoal or helminthic infections.				
XX					
SQ	Sequence 1224 BP; 284 A; 349 C; 297 G; 294 T; 0 U; 0 Other;				
	Query Match 57.8%; Score 947.8; DB 12; Length 1224;				

Db 787 TGGCTGCCCCGACTTGAAGATTTCCTCCATGGATGTCAGACATGTATCATGCAAC 846
 Qy 569 TTGAGAGCTCATCATCTCTGAGCCCTCTGCAATCTCTGTCACTTTGAGTGGGTACA 628
 Db 847 TGGAAAGCT-----TTGGATATA 864
 Qy 629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTGCTGTCAGTGGCTGAGG 688
 Db 865 CGATGAATGACCTCATCTTTGAGTGGCAGGACAGG---GAGCCGTGAGGTAGCAGATG 921
 Qy 689 GGCTGACTCTGCCCTGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTGTGACCAAGC 748
 Db 922 GACTAACTCTGCCCTGTTTATCTTGAAGGAGAGAGGACTTTGAGATCTGCAACCAAGC 981
 Qy 749 ACTACACACAGGGAATTCACCTGATCGAGTAAAGTTTACCTGGGAACGGCAGATGG 808
 Db 982 ACTACACACAGGTAATTCACCTGATTTGAGGCGCGGTTTCCACCTGGAGCGCAGATGG 1041
 Qy 809 GCTACTATCTGATTCAGATGTATACCTCCAGCCCTACTCATCGTCACTCTGCTGGGTCT 868
 Db 1042 GTTACTACCTGATTCAGATGTATTTCCAGCCCTGCTCATTTGCTCTCATGATCT 1101
 Qy 869 CTTCTGGATCAACATGGATGCTGCCCTGCTCCCGTGTGGGCTTGGGCATCACCACCGTGC 928
 Db 1102 CTTCTGGATCAACATGGATGCTGCCCTGCTCCCGTGTGGGCTTGGGCATCACCACCGTGC 1161
 Qy 929 TCACATGACACCCAGAGCTTGGCTCCGGGCTCTTTGCTTAAGTGTCTCTAGCTGA 988
 Db 1162 TCACATGACACCCAGAGCTCCGGGCTCTCGAGCATCTCTGCCAAGGTGCTCTATGTGA 1221
 Qy 989 AGCAATCGACATCTGAGTGTGCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
 Db 1222 AGCCATGACATTTGATGGCAGTTTGGCTGCTCTTTGGTGTCTCAGCCCTATTAGAAT 1281
 Qy 1049 ATGCTGCCATAAATTTGTTTCTCGTCAGATTAAGAAATTCATACGACTTCGAAGAAGC 1108
 Db 1282 ATGCTGCCATAAATTTGTTTCTCGTCAGATTAAGAAATTCATACGACTTCGAAGAAGC 1341
 Qy 1109 AGAGGCGCAACCGCTTGGAGGAGATATCATCAAGAAAGTCTGTTTCTATTTCCGTGGCT 1168
 Db 1342 GGAGACATCACAGGAGGATGAAGCT-----GGAGAAGGCGGCTTTAACTTCTGCTCT 1395
 Qy 1169 ATGGCTTGGGCT---ACTGCTCGCAGCAGAGATGGAGTCCCAATGGAGGTTCT---TG 1222
 Db 1396 ATGGATGGGCGCCAGCTGTCTACAGGCCAAGATGGCATCTCAGTCAAGGGCGCAACA 1455
 Qy 1223 GCATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGGAAGAGAAACCAACCGCAAAAC 1282
 Db 1456 ACAGTAACACCAACACCCCTCTTGCACCATCTAAGTCCCGAGGAGATGCGAAGAC 1515
 Qy 1283 TCTACGTGGACTGAGCCAGAGAAATGACACCATCTCCCGGGGTGTCTTCCCTTTCACTT 1342
 Db 1516 TCTTCATCCAGAGGCGCAAGATCGACAAAATATCCCGCATTTGGCTTCCCATGGCT 1575
 Qy 1343 TCTCATCTCATATCTTCTACTGGTGTCTATAAGTGTCTATGTCAGAGATATCC 1402
 Db 1576 TCTCATTTTCAACATGTTCTACTGGATCATCTACAAGATTGTCGCTAGAGAGAGCTCC 1635
 Qy 1403 ACCA 1406
 Db 1636 ACAG 1639

RESULT 11
 ADB85139
 ID ADB85139 standard; DNA; 3865 BP.
 XX
 AC ADB85139;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Rat neonatal glycine receptor NglyR gene SEQ ID NO:20.
 XX

KW rat; streptozocin; kinase; phosphatase; ion channel protein; receptor;
 KW transporter; G-protein coupled receptor; GPCR; DNA-binding proteins;
 XX protease; enzyme; analgesic; gene therapy; pain; diabetes; ds; gene.
 OS Rattus norvegicus.
 XX
 XX EP1284297-A2.
 XX 19-FEB-2003.
 XX 26-JUL-2002; 2002BP-00255228.
 XX 27-JUL-2001; 2001GB-00018354.
 PR 07-FEB-2002; 2002GB-00002880.
 XX
 XX (WARN) WARNER LAMBERT CO.
 XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
 WPI; 2003-364994/35.
 DR P-PSDB; ADB85138.
 XX
 XX Use of gene sequence that is down-regulated in response to streptozocin-
 PT induced diabetes, vector, host cell, animal, polypeptide and antibody, in
 PT screening of compounds for treating or diagnosing pain..
 XX
 PS Claim 1; Page 73-74; 256pp; English.
 CC The invention relates to a novel isolated gene sequence that is down-
 CC regulated in the spinal cord in response to streptozocin-induced
 CC diabetes, or comprising, hybridising or having at least 80% sequence
 CC identity to a sequence whose expression products are kinases,
 CC phosphatases, ion channel proteins, receptors, transporters, G-protein
 CC coupled receptor proteins, DNA-binding proteins, proteases or enzymes,
 CC given in the specification. A gene of the invention has analgesic
 CC activity, and may have a use in gene therapy. The gene sequences, vector,
 CC host cell, animal, polypeptide and antibody are useful for screening of
 CC compounds for diagnosing or treating pain. The kits are useful for
 CC simultaneous, separate or sequential detecting and/or quantifying down-
 CC regulation of a gene sequence in the spinal cord of a mammal in response
 CC to streptozocin-induced diabetes. The compound or pharmaceutical
 CC composition is useful as a medicament for treating or diagnosing pain.
 CC The present sequence represents a gene of the invention.
 XX
 SQ Sequence 3865 BP; 1087 A; 833 C; 889 G; 1056 T; 0 U; 0 Other;
 Query Match 39.9%; Score 654.8; DB 10; Length 3865;
 Best Local Similarity 70.8%; Pred. No. 1.5e-176;
 Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
 Qy 84 AAAAGAGGAGTCAATCTGGAAACCAAGGGTCCAGCCCATGTCCTCTGATTTCT 143
 Db 1256 AGACCATGACTCCAGGCTCTGGAACACATCCCTCGACAGCCCTGCTCTCAGATTTCTT 1315
 Qy 144 AGACAACCTTATGGGGCGAACAATCTGGATATGATGTCAGGATTCGGCCCAATTTAAAGG 203
 Db 1316 GGATAAATTAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTAAAGG 1375
 Qy 204 CCCACCGGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAACAGAC 263
 Db 1376 TCTCCAGTAAACGTTACTTGCATATTTTATCAACAGTTTGGATCGGTCAAGAAC 1435
 Qy 264 CACAATGGACTACCGGGTGAATGTCTTTGGGGCAACAGTGAATGACCCACGCTGTG 323
 Db 1436 CACCATGGACTACCGAGTGAACATTTTCTGACACAGCAGTGAACGATTCACGGCTGGC 1495
 Qy 324 CTACCGAGATATCTCTGATGACTCTCTGGACCTGATCCCTCCATGCTGCACTTCTG 383
 Db 1496 ATACAGTGTATCCAGATGATTTCCCTGGATTTGGATTCATCGATGCTGGATTTG 1555
 Qy 384 GAAGCAGACCTCTCTTTGCTAATGAGAAAGGGCCAACTTCATGAGTGACCCAGGA 443
 Db 1556 GAAACAGATTTGTTCTTTGCCAATGAGAAAGGACCAATTTCCATGATGTCACCACTGA 1615

444 CAACAAAGTTACTCGCATCTTCAAGAAATGGAAATGCTGTACAGCATCAGGCTGACCTT 503
 1616 TAACAAGTTGCTGGGATTTCCAAATATGGCAAGTCTCTACAGATTAGACTACCTT 1675
 504 CATTTTGTCTGCTGATGGAAGCTCAAGAACTTCCCATGGACATCCAGAGTGCAGAT 563
 1676 GACTTTATCTGTGCCATGGACCTCAAGAACTTCCAAATGGATGTCCAGACCTGTACAAT 1735
 564 GCAGCTTGAGACTCATCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTCAGTTGG 623
 1736 GCAGCTGAGAGTT-----TTGG 1753
 624 CTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTCTGTCTCAAGTGC 683
 1754 GTACACCATGAATGACTGATATTTGAGTGGTTAAGTATGCTC---CAGTACAGTTGC 1810
 684 TGAGGGCTGACTTGTGCCCCAGTTTATCTTGGGGATGAGAAGATCTAGGCTGTGTGAC 743
 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAAGAAAGAACTTTGGCTATTGACAC 1870
 744 CAAGCACTACAACACAGGAAATTCACCTGCATCGAGTTAAAGTTTCACTTGAACGCA 803
 1871 AAAGCATTACAACACTGGCAAGTTTACCTGCAATGAGGTCAAGTTTCACTTGAACGCA 1930
 804 GATGGCTACTTATCTGATTCAGATGTACATCCACGCTTACTCATCTCTCTCTG 863
 1931 GATGGCTATTATTGATTCAGATGTATATCCAGCTGCTGTATGATCATTTTGTCTG 1990
 864 GTCTCTCTTGTGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
 1991 GGTCTCTTTTGGATAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2050
 924 CGTCTCACCATCAACACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
 2051 CGTCTGCAATGACTACACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2110
 984 CGTGAAGGCAATGCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1043
 2111 TGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2170
 1044 GGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
 2171 GGAATATGAGCAGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2230
 1104 AAGCAGAGGCGCAAGCTTGGAGAGATATCATCCAGAGAGTCTGCTGCTGCTGCTG 1163
 2231 ACGACAGAGAGGCAAGTAAAGAGAGAGATGTTACTGCTGAAGCGCTTTTAACTTCAG 2290
 1164 TGGCTATGCTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223
 2291 CGGTTATGGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2347
 1224 CATTTATAGTCCCAACCTTCCAGCCCTCTTCTTAAGGGAAGGAGAAACCAACGCGAACT 1283
 2348 -----ACCTGCCAACCTTCCGCAACCCCAAGAGATGAGATGCTATCAAGAAGAA 2401
 1284 CTAGTGGAGTGAAGCAAGAGATTCAGACCATCTCCGGGCTGCTTCTTCTTCTTCTT 1343
 2402 GTTTGGATTCGGGCAAGAGATTCAGACCATCTTCTGAGTGCCTTCCCACTGCGCTT 2461
 1344 CTTCACTTCAATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1403
 2462 CCTCATTTCAACATCTTTTACTGGATCACAATCAAGATCATTCGCGCATGAAGATGCTCA 2521
 1404 CCAG 1407
 2522 CAAG 2525

RESULT 12
 AB218534
 _ID AB218534 standard; cDNA; 1857 BP.

XX AB218534;
 AC 23-JAN-2003 (first entry)
 DT Group III cDNA cancer related clone SEQ ID NO:960.
 DE Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.
 XX Homo sapiens.
 OS WO200278516-A2.
 PN 10-OCT-2002.
 PD 28-MAR-2002; 2002WO-US010421.
 PF 30-MAR-2001; 2001US-0280255P.
 PR 28-AUG-2001; 2001US-0315563P.
 PR 09-JAN-2002; 2002US-0347113P.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Wang S, Bangur CS, Gaiger A;
 PI WPI; 2003-058387/05.
 XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
 PT virology, immunology, microbiology, molecular biology and recombinant DNA
 PT techniques.
 XX Claim 1; SEQ ID NO 960; 207pp; English.
 PS ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines. (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques. N.B. The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1857 BP; 526 A; 416 C; 417 G; 498 T; 0 U; 0 Other;
 SQ
 Query Match 39.6%; Score 650.2; DB 8; Length 1857;
 Best Local Similarity 70.1%; Pred. No. 2.2e-175;
 Matches 943; Conservative 0; Mismatches 348; Indels 54; Gaps 3;
 QY 63 CTTCTCTCAGGGTGGCTTGGCAAAAGAGAAAGTCAAAATCTGGAACCAAGGGGTCCAGCC 122
 DB 450 CTTTCCAGGCGCTTCTGCAAGACCATGACTCCAGGCTCTGGAACCAACCTTTCACAGAC 509
 QY 123 CATGTCCCTCTGATTTCTTAGACAACTTATGGGCGAATCTGGATATGATGCCAG 182
 DB 510 CCTATCTCTCTCAGATTTCTTGGCAAGTAAATGGGAAGGACATCAGGATATGATCAAG 569
 QY 183 GATTGGGCGCAATTTTAAAGGCCCCCGTGAACGCTGACCTGCAACATCTTCAATCAACAG 242
 DB 570 AATCAGGCCAAATTTTAAAGGCTCTTCCAGTAAAGTACTTGCATATATTTTATCAACAG 629
 QY 243 TTTTCACTCCGTCAACAGACCAATGACTACCGGGTGAATGCTTCTTTCGGCAACA 302
 DB 630 TTTTGGATCAGTCACAGAAACGACATGACTACCGAGTGAATATTTTCTTGAGACACA 689
 QY 303 GTGGAATGACCAAGCTCTCTTACCGAATATCTGATGACTCTCTGGAACCTCATCC 362
 DB 690 GTGGAATGATTCACGGCTGGCTGACAGTACCGAGATGACTCCCTCGACTTGGACCC 749

QY 363 CTCCATGCTGGACTCTATCTGAGAGCCAGACCTCTTCTTGTCTATGAGAAAGGGCCAA 422
 Db 750 ATCCATGCTAGACTCCTATTTGGAAACCAAGATTTGTTCTTTGGCAATGAGAAGGGTGCCAA 809
 QY 423 CTTCCATGAGGTGACCAAGCAACCAAGTTACTGCGCATCTTCAAGAAATGGGAATGTCT 482
 Db 810 CTTCCAGATGTCACCACTGACACAAATTTGTACGGATTTGAAATAATGGCAAGTGT 869
 QY 483 GTACAGCATGAGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCAT 542
 Db 870 CTACAGTATCAGACTCACCTTGACCTTATCTGTGCTCCATGGACTTGAAGAACTTTCCGAT 929
 QY 543 GGACATCCAGAGTGCACCATGAGCTTGAAGCTCATCCATCTCTGAGCCCTCTGCC 602
 Db 930 GGATGCTCCAGACTGTACAAATGAGCTGGAGTT----- 964
 QY 603 ATCTCTGCTACCTTTTCACTTTGGCTTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGA 662
 Db 965 -----TTGGGTACACGATGATGACCTGATATTTGAGTGGTTAAGTGA 1007
 QY 663 TGTCTCTGCTGCTCAAGTGGCTGAGGGGTGACTCTGCCCCAGTTTATCTTGGGGATGA 722
 Db 1008 TGTCT---CAGTCAAGTTGCTGAAGCAATTGACCTGCCCCAGTTTATTTTGAAGAAGA 1064
 QY 723 GAAGATCTAGCTGTTGTACCAAGCACTACACACAGGGAATTCACCTGCTGATCGAGT 782
 Db 1065 GAAGGAATTTGGCTACTGTACAAAGCACTACACACTGGAAGATTTTACCTGCAATGAGT 1124
 QY 783 AAAGTTTACCTGGAAGCGCAGATGGCTACTATCTGATTCAGATGATACATCCCAAGCT 842
 Db 1125 CAAAGTTTACCTGGAAGCGCAATGGATATTTATTTGATCCAGATGATATCCCAAGCT 1184
 QY 843 ACTCATGCTGCTCTCTGCTGCTCTCTCTGATCAACATGATGCTGCCCCCTCCCG 902
 Db 1185 GCTATAGTAATTTTGTCTGCTGCTCTCTTGGATAAATGATGACGCCCCCTGCCAG 1244
 QY 903 TGTGGCTCTGGGATCACCACCGTGTCTACATGACCAACCGAGCTTGGCTTCCGGGC 962
 Db 1245 GGTGCACTGGGATCACCACAGTCTTAACGATGACCAACCGAGTTTCAAGGCTCCAGGC 1304
 QY 963 CTTCTTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 Db 1305 ATCTTCCCAAGGCTCTCTATGTAAGGATGATGATGATGATGATGATGATGATGAT 1364
 QY 1023 CTTTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
 Db 1365 GTTGTGTTGCTGCTTACTGGAATACGAGCGGTGACTTCTGCTCCAGGCAACAA 1424
 QY 1083 AGAATTCATACGATTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
 Db 1425 GGAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
 QY 1143 AGAAGTCTGTTTCTATTTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
 Db 1485 TGAAGTCTGTTTAAATTTTGGCGTTTATGGATGGGTCTGCTGCTGCTGCTGCTGCTGCT 1544
 QY 1203 AGTCCAAATGGAAGTTCTGGCATTTATATGATCCCAACCTCCAGCCCTCTTCTAAGGA 1262
 Db 1545 AACAGTGTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCCAAAGA 1595
 QY 1263 AGAGAAACACCGGAAACTCTAOTGATGAGCAAGAGAAATTTGACCACTTCTCCG 1322
 Db 1596 TGGAGATGCTATCAAGAAGAGTTTGTGAGCGGGCAAAAGGATTTGACACGATATCTCG 1655
 QY 1323 GGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1382
 Db 1656 AGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1715
 QY 1383 GCTATGCTCAGAGATATCCACAG 1407
 Db 1716 CATTCGGCATGAAGATGTCACAG 1740

RESULT 13
 ABA00851
 ID ABA00851 standard; DNA; 1386 BP.
 XX
 AC ABA00851;
 XX
 DT 16-APR-2003 (first entry)
 XX
 DE GLRA related nucleotide sequence #2.
 XX
 KW Glycine receptor subunit alpha; GLRA; p53 pathway; modulation; cancer;
 breast; colon; lung; ovary; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200299140-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 02-JUN-2002; 2002WO-US017458.
 XX
 PR 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Friedman L; Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 WPI; 2003-167349/16.
 XX
 PT Identifying p53 pathway modulating agents, useful for the diagnosis and
 treatment of disorders associated with defects in the p53 pathway, such
 as cancer, comprises assaying glycine receptor subunit alpha (GLRA)
 polypeptide or nucleic acid.
 XX
 PS Disclosure; Page 41-42; 56pp; English.
 XX
 CC The sequences given in ABA00850-55 represent glycine receptor subunit
 alpha (GLRA) related nucleotide sequences. These sequences may be used in
 the method of the invention for identifying a candidate p53 pathway
 modulating agent. The method comprises: (a) assaying a purified GLRA
 polypeptide or nucleic acid, or functionally active fragment or
 derivative, with a test agent; (b) detecting a test agent-biased activity
 of the assay system, and comparing the result with a reference. The
 method is useful for identifying a candidate p53 pathway modulating
 agent, modulating a p53 pathway of a cell and for diagnosing a disease in
 a patient. Diagnosing a disease in a patient, comprising obtaining a
 biological sample from the patient, contacting the sample with a probe
 for GLRA expression, comparing results with a control, and determining
 whether the results indicate a likelihood of disease. The disease is
 cancer of the breast, colon, lung or ovary having a greater than 25%
 expression level. The methods and compositions of the present invention
 are useful for the diagnosis and treatment of disorders associated with
 defects in the p53 pathway, such as cancer of the breast, colon, lung and
 ovary
 XX
 SQ Sequence 1386 BP; 393 A; 320 C; 317 G; 356 T; 0 U; 0 Other;
 Query Match 39.5%; Score 648.6; DB 10; Length 1386;
 Best Local Similarity 70.0%; Pred. No. 5.3e-175;
 Matches 942; Conservative 0; Mismatches 349; Indels 54; Gaps 3;
 QY 63 CCTCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAATCTGGAACCAAGGGTCCAGCC 122
 Db 68 CTTCAGGAGCGCTTCTTGCAAAAGACCATCTCCAGGTCTGGAACAACACCTTCACAGAC 127
 QY 123 CATGTCCCCTCTGATTTCTAGACAAACTTATGGGGCGCAACATCTGGATATGATGCCAG 182
 Db 128 CCTATCTCTCTCAGATTTCTTGACAAAGTTTATGGGAAGACATCAGGATATGATCGAAG 187
 QY 183 GATTCGGCCCAATTTTAAAGGCCCCACCCCGTGAACGTGACCTGCAACATCTTTCATCAACAG 242

188 AATCAGGCGCAAAATTTTAAAGGCTCTCCAGTAACAGTGTACTTGCAATATTTTATCAACAG 247
 243 TTTTACGCTCGGTCAACCAAGACCAATGAGTACCGGGTGAATGTCTTCTTGGGGCAACA 302
 248 TTTTGGATCAGTCACAGAAGCAACGATGAGTACCGAGTGAATATTTTCTGAGACACA 307
 303 GTGGAATACCCAGCGCTGTCTTACCGAGAATATCTGTGATGATCTCTGTGACCTCGATCC 362
 308 GTGGAATGATTACAGCGCTGCGGTACAGTGAGTACCCAGATGACTCTCTGGACTTGGACCC 367
 363 CTTCATGCTGGACTCTATCTGAGAGCGACACTCTTCTTTTGTATGAGAAAGGGGCCAA 422
 368 ATCCATGCTAGACTCTCAATTTGGAACACAGATTGTCTTTTGGCAATGAGAAGGGTCCCAA 427
 423 CTTTCCATGAGGTGACCAACGACCAACAAAGTTACTGCGCATCTTCAAGAAATGGGAATGTGCT 482
 428 CTTTCCAGATGTCAACCACTGACACAAATTTGCTACGGAATTCGAAATATGGCAAGTGTCT 487
 483 GTACAGCATCAGGCTGACCTCATTTTGTCTGCTGTGATGAGCTCAAGAACTTCCCCAT 542
 488 CTACAGTATCAGACTCACCTTGACCTTATCTGTCCATGGACTTGAAGACTTTCCGAT 547
 543 GGACATCCAGACGTGACAGATGAGCTTGAAGCTCATCCATCTCTGAGCCCTCTGCC 602
 548 GGATGTCCAGACCTGTACAAATGACGTGGAGATT----- 582
 603 ATCTCTGTCACTTTCATTTGGTGGCTACACCATGAAGACCTCGTGTGTGAGTGGCTGGAACA 662
 583 -----TTGGGTACAGATGATGACCTGATATTTGATGGTGTAAAGTGA 625
 663 TGCTCTGTCTGCTCAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGATCA 722
 626 TGGTCT---CAGTCAAGTTGCTGAAGATTGACCTGTCCCTGCCCCAGTTTATTTGAAGAAGA 682
 723 GAAGGATCTAGGCTGTGTACCAAGCACTACAAACAGGGAAATTCACCTGATCATGAGT 782
 683 GAAGGAATCTGGTACTGTACAAAGCACTACAACTGGAAGATTACCTGATGAGT 742
 783 AAAGTTTCACTTGAAGCGCAGATGGCTACTATCTGATTCAGATGTACATCCCGACCT 842
 743 CAAAGTTTCACTTGAAGCGCAATGGATATTTATTTGATCCAGATGATATCCCAAGCT 802
 843 ACTCATGCTGATCTCTCTGCTGCTCTCTTCTGATCAACATGATGCTGCCCCCTCGCG 902
 803 GCTATAGTAATTTTCTGCTGGTCTCTTTTGGATAAATATGATGACAGCCCTGCGCAG 862
 903 TGTGGGCTTGGGATCACACCGTGTCTACCATGACACCCAGAGCTCTGGCTCCCGGC 962
 863 GGTGCGACTGGGCATCACACAGTCTTAACGATGACACCCAGAGTTCAAGGCTCCAGGGC 922
 963 CTCTTGTGCTTAAAGTCTCTGCTGAGGCAATCGACATCTGATGGCTGTGTGCTGCT 1022
 923 ATCTCTGCGCAAGGCTCTCTATGTAAGCGATTTGATCTGATGGCGGTGTGCTTCT 982
 1023 CTTTGTGTTGCTGCTGCTGCTGATGCTGCTCAATTTTGTCTGCTGAGATATAA 1082
 983 GTTGTGTTGCTGCTTACTGGAATACGAGGGTGAATCTGCTTCCAGGCAACACAA 1042
 1083 AGAATTCATACGATTTGGAAGAGCAGAGCGGCCAACCGCTTGGAGGAAGATATCATCA 1142
 1043 GGAAGTCTCTGCGCTCCGAAGAAGACAGAGAGGCGAGATAAGGAAGAAGAGCTTACTCG 1102
 1143 AGAAGTCTCTTATTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
 1103 TGAAGTCTGTTTAAATTTAGCGGTTATGGGATGGGTCTGCTGCTGCTGCTGCTGCTGCT 1162
 1203 AGGTCCATGGAAGGTTCTGGCATTTATATGATCTCCCAACCTCCAGCCCTCTTCTTAAGGA 1262
 1163 AACAGCTGTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCCAAAGA 1213
 1263 AGAGAAACACCGGAAATCTTACGTGAGCTGAGCGCAAGAAATTTGACCAATCTCCCG 1322

Db 1214 TGGAGATGCTATCAGAAGAGAGTTTGTGGACCGGGCAAAAGGATTGACACGATATCTCG 1273
 QY 1323 GGCTGTCTTCCCTTTTCACTTCTCTCATCTTCAATATCTTCTACTGGTGTCTATAAAGT 1382
 Db 1274 AGCTGCTTCCCATTTGGCTTCTCTCATTTTCAACATCTTTTACTGGATCACATACAGAT 1333
 QY 1383 GCTATGGTCAGAAGATATCCACCAG 1407
 Db 1334 CATTCGGCATGAAGATGTCACAAG 1358
 RESULT 14
 ABA00850
 ID ABA00850 standard; DNA; 1854 BP.
 XX ABA00850;
 AC ABA00850;
 XX 16-APR-2003 (first entry)
 DT GLRA related nucleotide sequence #1.
 DE Glycine receptor subunit alpha; GLRA; p53 pathway; modulation; cancer;
 KW breast; colon; lung; ovary; ss.
 KW Homo sapiens.
 OS WO200299140-A1.
 XX 12-DEC-2002.
 PD 02-JUN-2002; 2002WO-US017458.
 XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0352553P.
 XX (EXEL-) EXELIXIS INC.
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX WPI; 2003-167349/16.
 DR Identifying p53 pathway modulating agents, useful for the diagnosis and
 PT treatment of disorders associated with defects in the p53 pathway, such
 PT as cancer, comprises assaying glycine receptor subunit alpha (GLRA)
 PT polypeptide or nucleic acid.
 XX Example; Page 40-41; 56pp; English.
 PS The sequences given in ABA00850-55 represent glycine receptor subunit
 CC alpha (GLRA) related nucleotide sequences. These sequences may be used in
 CC the method of the invention for identifying a candidate p53 pathway
 CC modulating agent. The method comprises: (a) assaying a purified GLRA
 CC polypeptide or nucleic acid, or functionally active fragment or
 CC derivative, with a test agent; (b) detecting a test agent-biased activity
 CC of the assay system, and comparing the result with a reference. The
 CC method is useful for identifying a candidate p53 pathway modulating
 CC agent, modulating a p53 pathway of a cell and for diagnosing a disease in
 CC a patient. Diagnosing a disease in a patient, comprising obtaining a
 CC biological sample from the patient, contacting the sample with a probe
 CC for GLRA expression, comparing results with a control, and determining
 CC whether the results indicate a likelihood of disease. The disease is
 CC cancer of the breast, colon, lung or ovary having a greater than 25%
 CC expression level. The methods and compositions of the present invention
 CC are useful for the diagnosis and treatment of disorders associated with
 CC defects in the p53 pathway, such as cancer of the breast, colon, lung and
 CC ovary
 XX Sequence 1854 BP; 525 A; 415 C; 416 G; 498 T; 0 U; 0 Other;
 SQ Query Match 39.5%; Score 648.6; DB 10; Length 1854;
 Best Local Similarity 70.0%; Pred. No. 6.2e-175;
 Matches 942; Conservative 0; Mismatches 349; Indels 54; Gaps 3;

QY 63 CCTCCTCAGGTGGCTTTCGCAAGAGAGAGTCAATCTGAAACCAAGGGGTCCAGCC 122
Db 447 CTTACAGGCGGCTTTCGCAAGAGACATGACTCCAGGTCTGGAACCAACCTTTCAGAC 506
QY 123 CATGTCCCTCTGATTTCTTAGACAACTTATGGGCGAAATCTGGATATGATGCCAG 182
Db 507 CTAICTCTCTCAGATTTCTTGACAAAGTAAATGGGAAGACATCAGGATATGATGCAAG 566
QY 183 GATTCGGCCCAATTTTAAAGGCCCAACCCGTGAAACGTGACCTGCAACATCTTCATCAACAG 242
Db 567 AATCAGGCCAAATTTTAAAGGCTCTCCAGTAAAGGTTACTTGCATATTTTATCAACAG 626
QY 243 TTTACGTCTGTCACCAAGACCAATGACTACCGGTGAATGCTTCTTGGGGAACA 302
Db 627 TTTTGGATCAGTCACAGAAACCAATGACTACCGAGTGAATATTTTCTGAGACAACA 686
QY 303 GTTGAATGACCAAGCGCTCTCTACCGAATATCTGATGACTCTCTGCACTGATCC 362
Db 687 GTTGAATGATTCAGGCTGGCTGACAGTACCGAGTACCGAGTACTCCCTGGACTTGGACCC 746
QY 363 CTCATGTCTGACTCTATCTGGAAGCCAGACCTCTTCTTTGCTAATGAAAGGGGCCAA 422
Db 747 ATCCATGCTAGACTCCATTTGGAACACAGATTGTTCTTTGCCAATGAGAAGGTGCCAA 806
QY 423 CTTCCATGAGGTGACCAACGACCAACAAAGTTACTGCGCATCTTCAAGAAATGGGAATGTCT 482
Db 807 CTTCCACGATGTCACCACTGACAAACAAATTTGCTACGGATTTGGAATAATGGCAAGTCT 866
QY 483 GTACAGCATCAGCTGACCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCAT 542
Db 867 CTACAGATCAGACTCACCTTGACCTTATCTCTGCTCCATGGACTTGAAGAACTTTCCGAT 926
QY 543 GGACATCCAGAGCTGACGATGACGCTGAGAGCTCATCCACTCTGCAGCCCTCTGCC 602
Db 927 GGATGTCCAGACTGTACATGACGCTGGAGTT----- 961
QY 603 ATCTCTGTCACTTTAGTTGGCTGTACACCATGAAAGACCTCGTGTGATGCTGGAAGA 662
Db 962 -----TTGGGTACACGATGAATGATGATCTGATTTTGAAGTGA 1004
QY 663 TGCTCTGCTGCTCAAGTCTGAGGGCTGACTCTGCCCGAGTTTATCTTGGGGATCA 722
Db 1005 TGGTC---CAGTGAAGTGTCTGAAGGATGACCTCGCCCGAGTTTATTTGAAAGA 1061
QY 723 GAAGGATCTAGGTGTGTACCAAGCACTACACACAGGGAATTCACCTGATCGAGT 782
Db 1062 GAAGGAACTTGGCTACTGTACAAAGCACTACACACTGGAAGTTTACCTGCAATTGAGT 1121
QY 783 AAAGTTTCACTGGAAACGCGAGATGGGCTACTATCTGATTCAGATGATATCCCAAGCT 842
Db 1122 CAAAGTTTCACTGGAAACGCGCAATGGATATTTATTTGATCCAGATGTATCCCAAGCT 1181
QY 843 ACTCATGCTCATCTCTCTGGTCTCTCTGATCAACATGGATGCTCCCTGCGCG 902
Db 1182 GCTTATAGTAATTTCTCTGGGTTTCTTTTGGATAAATATGATGCAAGCCCTGCCAG 1241
QY 903 TGTGGCCCTGGGATCACCACCGTGTCTACCATGACACCCAGAGCTCTGGCTCCCGGC 962
Db 1242 GGTGCACTTGGGATCACCACAGTCTTAAGATGACCAACAGAGTTGAGGCTCAGGGC 1301
QY 963 CTCCTTTGCCCTAAGGTCTCTACGTGAAGCAATCGACATCTGGATGGCTGTGTCTGCT 1022
Db 1302 ATCTCTGCCAAAGGTCTCTATGTAAGAGCATTTGATGATGCGGTGTGCTTCT 1361
QY 1023 CTTTGTGCTGCTGCTGCTGAGATGCTGCAATTAATTTTGTCTGCTGAGATTA 1082
Db 1362 GTTTGTGTTGCTGCTTACTGGAATACGAGCGGTGAACTTCTGCTCCAGGCAACAA 1421
QY 1083 AGAATTCATACGATTCGAAAGAGGAGGAGGCGCCCAAGCTTGGAGGAAGATATATCA 1142
Db 1422 GGAGTTCTCTGCGCTCCGAAAGAGACAGAGGCAAGATGAAGGAAGAGACGTTACTCG 1481

QY 1143 AGAAGTCGTTTCTATTTCCGTGGCTATGGCTTTGGSCCACTGCCTGACGCAAGAGATGG 1202
Db 1482 TGAAGTCGTTTAAATTTTAGCGGTTATGGGATGGTCACTGCTCCAAAGTGAAGATGG 1541
QY 1203 AGGTCCAATGGAAGGTTCTGGCATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGA 1262
Db 1542 AACAGCTGTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCAAAAGA 1592
QY 1263 AGGAGAAACCAACGCGGAACACTCTAGCTGAGCTGAGCAAGAGAAATTTGACACCATCTCCCG 1322
Db 1593 TGGAGATGCTATCAAGAAGAAAGTTTGTGACCGGGCAAAAGGATTTGACACGATATCTCG 1652
QY 1323 GCCTGCTTCTCCCTTTTCACTTCTCTCATCTTCAATATCTTCTACTGGTGTCTATAAAGT 1382
Db 1653 AGCTGCTTCTCCCATTTGGCCTTCTCATTTTCAACATCTTTTACTGGATCACATACAGAT 1712
QY 1383 GCTATGGTCAGAAGATATCCACAG 1407
Db 1713 CATTGGCATGAAGATGTCCACAAG 1737

RESULT 15
ADC77686

ID ADC77686 standard; cDNA; 3069 BP.

XX AC ADC77686;

XX XX 01-JAN-2004 (first entry)

XX DT Human 2387 protein encoding cDNA SEQ ID NO:49.

XX DE pain disorder; pain signalling mechanism; analgesic; antimigraine;
XX KW antinflammatory; gene therapy; inflammatory pain; chronic pain;
XX KW neuropathic pain; neuralgia; fibromyalgia; cancer pain; migraine;
XX KW headache; pain; human; gene; ss.

OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 421..1770

FT /*tag= a

FT /product= "2387 protein"

XX PN WO2003073983-A2.

XX 12-SEP-2003.

XX 19-FEB-2003; 2003WO-US004816.

XX 28-FEB-2002; 2002US-0360495P.

PR 04-APR-2002; 2002US-0370121P.

PR 16-APR-2002; 2002US-0373010P.

PR 19-APR-2002; 2002US-0373908P.

PR 03-MAY-2002; 2002US-0377717P.

PR 13-MAY-2002; 2002US-0379949P.

PR 03-JUN-2002; 2002US-0382409P.

PR 06-JUN-2002; 2002US-0385280P.

PR 10-JUN-2002; 2002US-0387536P.

PR 08-JUL-2002; 2002US-0394376P.

PR 21-AUG-2002; 2002US-0404996P.

PR 19-SEP-2002; 2002US-0412006P.

PR 09-OCT-2002; 2002US-0417327P.

PR 10-OCT-2002; 2002US-0417499P.

PR 15-NOV-2002; 2002US-0426964P.

PR 10-DEC-2002; 2002US-0432320P.

XX XX (MILL-) MILLENNIUM PHARM INC.

XX PI Rosenfeld JB, Silos-Santiago I;

XX WPI; 2003-712843/67.

DR P-PSDB; ADC77687.

XX Identifying a compound capable of treating a pain disorder e.g.,
PT neuropathic pain comprises assaying the ability of the compound to
PT modulate the nucleic acid expression or polypeptide activity.
XX
PS Claim 1; SEQ ID NO 49; 277bp; English.
XX
CC The present invention describes a method for identifying a compound (C)
CC capable of treating a pain disorder comprising assaying the ability of
CC the compound to modulate 9949, 14230, 760, 62553, 12216, 17719, 41897,
CC 47174, 33408, 10002, 16209, 314, 636, 27410, 33260, 619, or 13424 nucleic
CC acid expression or 9949, 14230, 760, 62553, 12216, 17719, 41897, 47174,
CC 33408, 10002, 16209, 314, 636, 27410, 33260, 619, 15985, polypeptide
CC activity. Also described: (1) identifying a compound (C) capable of
CC modulating a pain signalling mechanism; and (2) treating a subject having
CC a pain disorder characterised by aberrant nucleic acid expression or
CC polypeptide activity. (C) has analgesic, antimigraine and
CC antiinflammatory activities, and can be used in gene therapy. The method
CC is useful for identifying a modulator compound capable of treating a pain
CC disorder, e.g. inflammatory pain, chronic pain, neuropathic pain,
CC neuralgia, fibromyalgia, cancer pain, migraine/headache pain or tissue
CC pain comprising administering the modulator to a subject having a pain
CC disorder characterised by aberrant nucleic acid expression or polypeptide
CC activity. The present sequence encodes the human 2387 protein from the
CC present invention.
XX
XX Sequence 3069 BP; 962 A; 547 C; 583 G; 977 T; 0 U; 0 Other;
Query Match 33.7%; Score 552.8; DB 10; Length 3069;
Best Local Similarity 66.6%; Pred. No. 2.5e-147;
Matches 885; Conservative 0; Mismatches 377; Indels 66; Gaps 4;
QY 80 TGGCAAGAGGAGTCAATCTGGAAACCAAGGGGTCCCGCCATGTCCCTCTGATT 139
DB 497 TTGCCCAAGAAACAGACAGTGCAGATCTCGAAGTCTCGAATGTCTCAATCTCTGATT 556
QY 140 TCCTAGACAACTTATGGGGCGAACAATCTGGATATGATCCAGGATTCGGCCCAATTTTA 199
DB 557 TTCTGGATAATTAATGGGCGAGACATCAGGATATGATGCAAGATCAGACCAATTTTA 616
QY 200 AGGCCCAACCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTTCAGCTCGGTCAACA 259
DB 617 AAGGCCCTCCAGTTAATGTGCATGCAACATATTCATCAACAGTTTCGGCTCTATCGCAG 676
QY 260 AGACCAATGACTACCGGTGAATGTCTTCTGGGCAACAGTGGATGACCCACGCC 319
DB 677 AGACGACCATGGATTACAGATGAAATCTTCTTCGTCAGAAATGGAATGATCCCCGCC 736
QY 320 TGTCTACCGAGAAATATCTGTGACTCTCTGGACCTCGATCCCTCCATCTGCGACTTA 379
DB 737 TCGCGTACAGTGAATATCTCGAGACTCTTTAGACCTCGACCCCTCCATGTTGGACTCCA 796
QY 380 TCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGCCAACTTCCATGAGGTGACCA 439
DB 797 TTGGAAACCTGATTGTTCTTTGGCCAAATGAAAGGGTGCACACTTTCATGAAGTCACTA 856
QY 440 CGGACAAACAAAGTTACTCGGCATCTTCAAGATGGGAATGTGTGTACAGCATCAGGCTGA 499
DB 857 CAGACAAACAAATGCTAAGAAATTTTCAAAATGGAATGTTCTTTAATCAATAAGATTAA 916
QY 500 CCCTCAATTTGTCTGCTGATGAGCTCAAGAACTTCCCATGGACATCCAGACGTGCA 559
DB 917 CATTAACACTTTCTCTGCCAATGATCTCAAGAAATTTTCCCATGGATGTACAAACATGTA 976
QY 560 CGATGAGCTTGAAGCTCATCATCTCTGCGACCTCTGCGACCTCTGCTACCTTTCAG 619
DB 977 TAATGCAACTGGAAAGCT----- 994
QY 620 TTGGCTTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTGCTGTCCAAAG 679
DB 995 TTGGGTACACATGAATGATCTCATTTTGAATGGC---AAGATAGGCCACCCGTTACAAG 1051
QY 680 TGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTT 739

1052 TGGCAGAGGAGTCACTCTTTGCCCCAGTTTCTCTGTAAGAGAGAAAGATTACGATACT 1111
QY 740 GTACCAAGCACTACACACAGAGGAAATTCACCTGCATCGAGGTAAAGTTTTCACCTGGAAC 799
DB 1112 GCACATAACATTACAAATACAGGAAAGTTTACGTGTATAGAGTGGGATTCCCATCTGGAGC 1171
QY 800 GGCAGATGGGCTACTATCTGATTTCAGATGTACATCCCGAGCCTACTCATCGTCACTCTGT 859
DB 1172 GACAAATGGGATACTATCTGATCCAGATGTACATTCACAGTCTCTCTGATGTTTATCTAT 1231
QY 860 CTTGGGTCTCTCTCTGATCAACATGGATGTCGCCCTGCGCTGCGCTGCGCTGCGCTGCGCT 919
DB 1232 CTTGGGTCTCTCTCTGATCAACATGGATGTCAGCAGCCGCGCAGGTAGCTCTGGGATAA 1291
QY 920 CCACGCTGCTACCATGACACCCAGAGCTCTGGCTCCCGGCGCTCTTTCCTCAAGGTGT 979
DB 1292 CCACCGCTGCTAAGATGACTACACAGATTCAGGATTCAGGCTTCCTTTCGCAAAAGTTT 1351
QY 980 CTTAGCTGAAGCAATCGACATCTGATGGCTGTGTCTGTCTCTTGTGTTGCTGCTGCT 1039
DB 1352 CATATGTCAAAGCTATGTATTTGGATGGCAGTATGCTCTCTTGTGTTGTTTTCAGCAC 1411
QY 1040 TCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
DB 1412 TTTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471
QY 1100 GAAGAGGAGGAGGCGCCAAAGCTTGGAGGAGATATATCCCAAGAAAGTCTGTTCTATT 1159
DB 1472 GACGAAAGAGAAAG-----AATAAGGATGATGAGGTAAAGGAAAGCCGATTCAGCT 1522
QY 1160 TCGGTGGCTATGGCTTGGGCCACTGCTGCGAGCAAGAGATGGAGTCCAAATGGAGGTT 1219
DB 1523 TCACAGCTTATGGATGGAGCCATGCTTACAGCAAGAGATGGCATGACTCCAAAGGGCC 1582
QY 1220 CTGGCAATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGAAAGAGAGAAACCAAGCGGA 1279
DB 1583 C-----CAACCACTGCTCCAGGTAAATGCCAAAGAGTCTGATGAAATGAGGA 1630
QY 1280 AACTCTACGTGACCTGAGCCAAAGAAATTCACCATCTCCCGGCTGTCTTCCCTTTCA 1339
DB 1631 AGGTCTTTATCGACCGGCGCAAGAAATGATACCATCTCCCGAGCTCTTCCCAATTAG 1690
QY 1340 CTTTCTCATCTTCAATATCTTCTACTGGTGTGTCTATAAGTGTCTATGGTCAGAGATA 1399
DB 1691 CTTTGTGATTTTATATTTTCTACTGGTTATCTATATAAATTTCTTAGGCATGAGATA 1750
QY 1400 TCCACGAG 1407
DB 1751 TTCTATCAG 1758

Search completed: April 28, 2005, 12:05:58
Job time : 1101.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 11:47:18 ; Search time 337.133 Seconds
(without alignments)
7959.755 Million cell updates/sec

Title: US-10-075-846-3
Perfect score: 1640
Sequence: 1 agacaactctgtcttcgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	656	40.0	1372	4	US-09-949-016-4602 Sequence 4602, Ap
2	656	40.0	1715	4	US-09-949-016-27 Sequence 27, Appl
3	648.6	39.5	1854	4	US-09-949-016-5267 Sequence 5267, Ap
4	253	15.4	253	4	US-09-016-434-271 Sequence 271, Appl
5	211	12.9	2109	4	US-09-949-016-3698 Sequence 3698, Ap
6	193	11.8	1555	3	US-08-809-802-11 Sequence 11, Appl
7	187	11.4	4621	3	US-09-592-891A-13 Sequence 13, Appl
8	187	11.4	4621	4	US-09-969-844-13 Sequence 13, Appl
9	182.2	11.1	2310	1	US-08-417-330A-15 Sequence 15, Appl
10	181.6	11.1	1609	3	US-09-592-891A-11 Sequence 11, Appl
11	181.6	11.1	1609	4	US-09-969-844-11 Sequence 11, Appl
12	179	10.9	3958	1	US-08-435-933-5 Sequence 5, Appl
13	179	10.9	3958	5	PCT-US96-06035-5 Sequence 5, Appl
14	176.8	10.8	73519	4	US-09-949-016-16344 Sequence 16344, A
15	176.8	10.8	105919	4	US-09-949-016-11769 Sequence 11769, A
16	174.2	10.6	1640	3	US-09-592-891A-12 Sequence 12, Appl
17	174.2	10.6	1640	4	US-09-969-844-12 Sequence 12, Appl
18	168	10.2	1368	3	US-09-130-339-1 Sequence 1, Appl
19	161.2	9.8	1297	3	US-09-130-339-3 Sequence 3, Appl
20	158.8	9.7	1884	1	US-09-403-667A-3 Sequence 3, Appl
21	157.8	9.6	1657	1	US-08-554-659-3 Sequence 3, Appl
22	157.8	9.6	1844	3	US-09-002-361-1 Sequence 1, Appl
23	157.2	9.6	1884	4	US-09-403-667A-1 Sequence 1, Appl
24	156.4	9.5	1866	1	US-08-417-330A-19 Sequence 19, Appl
25	156.2	9.5	1657	1	US-08-554-659-1 Sequence 1, Appl
26	156.2	9.5	1732	1	US-08-417-330A-17 Sequence 17, Appl
27	155	9.5	205163	4	US-09-949-016-17009 Sequence 17009, A

28	146.6	8.9	669	3	US-09-002-361-7 Sequence 7, Appl
29	143.2	8.7	1997	4	US-09-620-312D-256 Sequence 256, App
30	141.8	8.6	1491	1	US-08-137-614A-3 Sequence 3, Appl
31	141.8	8.6	1491	2	US-08-768-301-1 Sequence 1, Appl
32	138.6	8.5	1638	1	US-08-417-330A-13 Sequence 13, Appl
33	137.6	8.4	1519	3	US-09-002-361-4 Sequence 4, Appl
34	136.2	8.3	1038	4	US-09-765-069-9 Sequence 9, Appl
35	136.2	8.3	1179	4	US-09-765-069-3 Sequence 3, Appl
36	136.2	8.3	1263	4	US-09-765-069-7 Sequence 7, Appl
37	136.2	8.3	1404	4	US-09-765-069-1 Sequence 1, Appl
38	136.2	8.3	1987	4	US-09-620-312D-331 Sequence 331, App
39	135.6	8.3	2191	1	US-08-417-330A-11 Sequence 11, Appl
40	135.2	8.2	1052	3	US-09-592-891A-8 Sequence 8, Appl
41	135.2	8.2	1052	4	US-09-969-844-8 Sequence 8, Appl
42	131.6	8.0	2066	3	US-08-072-064-7 Sequence 7, Appl
43	130.6	8.0	1693	4	US-09-620-312D-131 Sequence 131, App
44	130.6	7.9	2066	3	US-08-072-064-2 Sequence 2, Appl
45	128.8	7.9	1679	4	US-09-016-434-1470 Sequence 1470, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-4602
; Sequence 4602, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4602
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4602

Query Match		40.0%;	Score 656;	DB 4;	Length 1372;
Best Local Similarity		71.9%;	Pred. No. 1.1e-186;		
Matches		952;	Conservative 0;	Mismatches 315;	Indels 57; Gaps 5;
QY	89	AGGAAGTCAAAATCTGGAAACAAGGGGTCCAGCCCATGTCCCTCTGATTTCCTAGACA	148		
Db	26	AGAGGCTGAGCTGCTCGTCCGACCAAGCCTATGTACCTCGGATTCCTGGATA	85		
QY	149	AACTTATGGGGCGAACAATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC	208		
Db	86	AGCTAATGGGAGAACCTCCGGATATGATGCCAGGATGAGGCCCAATTTTAAAGGTCCTCC	145		
QY	209	CGTGAACGTGACCTGCAACATCTTCATCAACAGTTTACGCTCCGTCACCAAGACCACAA	268		
Db	146	CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGGTTCCATTTGCTGAGAACCA	205		
QY	269	TGAGCTACCGGGTGAATGCTTCTTCGGGCAACAGTGAATGACCCAGCGCTGCTCTTACC	328		
Db	206	TGAGCTATAGGTCAACATCTTCTCGGAGCAATGGAACGACCCCGCTGGCTTATTA	265		
QY	329	GAGATATCTCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATCTGAAGC	388		
Db	266	ATGAATACCTGACGACTCTCTGGACCTGAGCCCACTCCATGCTGGACTCCATCTGGAAC	325		
QY	389	CAGACCTCTTTTGTGTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCCAGGACACA	448		

Db 326 CTGACCTGTTCTTGGCAAGAGAGGGGCCACCTTCATGAGATCAACACAGACAAC 385
 Qy 449 AGTTACTGCGCATCTTCAAGATGGGAATGTGCTGTACAGCATCAGGCTGACCTCATTT 508
 Db 386 AATTGCTAAGGATCTCCCGGAATGGGAATGTCTCTACAGCATCAAGATCAACCTGACAC 445
 Qy 509 TGTCTGCTGTGAGACCTCAAGAACTTCCCATGAGATCCAGAGCTGACAGATGACG 568
 Db 446 TGGCTTGGCCATGAGCTTGAAGATTTCCCATGAGATGTCAGACATGATATGCAAC 505
 Qy 569 TTGAGAGCTCATCATACTCTGAGCCCTCTGCCATCTCTGTCACTTTTCAGTTGGCTACA 628
 Db 506 TGGAAAGCT-----TTGGATATA 523
 Qy 629 CCATGAAGACCTCGTTTGTAGTGGCTGGAAGATGCTCTGCTGCTCAAGTGGCTGAGG 588
 Db 524 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCCGTGACAGGATGAG 580
 Qy 689 GGTGACTCTGCCCAAGTTTATCTTCGGGATGAGAGGATCTAGGCTGTGTACCAAGC 748
 Db 581 GACTAATCTGCCCAAGTTTATCTTGAAGAGAGAGGACTTGATGATGATGATGATGATG 640
 Qy 749 ACTACAACACAGGGAATTCACCTGCATCGAGTAAAGTTTCACTTGGAAACGGCAGATGG 808
 Db 641 ACTACAACACAGGTAATTCACCTGCATTTGAGCCCGGTTCCACCTGGAGCGCAGATGG 700
 Qy 809 GCTACTATCTGATGAGATGATATCCCGAGCTTACTCATCTGCTATCTCTGCTGCTGCT 868
 Db 701 GTTACTACCTGATGATGATATTTCCAGCGCTGCTCATTTGCTATCTCTCATGATCT 760
 Qy 869 CTTCTGGATCAACATGGATGCTGCCCTGCCGTGGGCTGGGCTGAGCATCAACACCGTCG 928
 Db 761 CTTCTGGATCAACATGGATGCTGACCTGCTGCTGCTGGGCTGAGCATCAACCTGCTG 820
 Qy 929 TCACATGACCAACAGAGCTCTGGCTCCCGGCTCTTTTGGCTAAGGTGCTCTACGTGA 988
 Db 821 TCACATGACCAACAGAGCTCTGGCTCTCGAGCATCTCTGCCCAAGGTGCTCTATGTC 880
 Qy 989 AGGCAATGACATCTGGATGGCTGTGTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTG 1048
 Db 881 AAGCCATTGACATTTGGATGGCAGTTTGGCTGCTCTTGTGCTCTCAGGCCCTATTAGA 940
 Qy 1049 ATGCTGCCATTAATTTTGTCTGCTGAGTAAAGAAATTCATACGACTTCGAAGAGGC 1108
 Db 941 ATGCTGCCGTTAATTTGTCTGCTGCAACATTAAGAGCTGCTCCGATTCAGAGGAAGC 1000
 Qy 1109 AGAGCGCCAAACGCTTGGAGGAAGATPATCATCAAGAAAGTCTTTTCTATTTCCGTGGCT 1168
 Db 1001 GGAGACATCAAGAGGAGTGAAGCT-----GGAGAAGCGCGCTTTAATCTCTGCTG 1054
 Qy 1169 ATGGCTTGGCC---ACTGCTGCGAGGAGAGATGAGGTCCATGGAAGGTTCT---TG 1222
 Db 1055 ATGGGATGGGCCAGCGCTGTCTACAGGCCAAGGATGGCATCTCAGTCAAGGGCGCCAA 1114
 Qy 1223 GCATTTATAGTCCCAACCTTCAAGCCCTCTTCTAAGGGAAGAGAAACACAGCGGAAC 1282
 Db 1115 ACATGAACCAACACCCCTCTCTGACATCTAAGTCCCAAGAGGAGATGCGAAGAC 1174
 Qy 1283 TCTACGTGGAATGAGCAAGAAATTTGACACCATCTCCCGGGCTGTCTTCCCTTTCAC 1342
 Db 1175 TCTTATCATGAGGCGCAAGAGATGCAACAAATATCCCGCATTTGCTTCCCATGGCT 1234
 Qy 1343 TCTCATCTTCAATATCTTCTACTGGGTGTCTATAAAGTGTATGGTCAGAGATATCC 1402
 Db 1235 TCTCATTTTCAATATCTTCTACTGGATCATCTACAAGATTTGTCGCTAGAGAGGAGCTCC 1294
 Qy 1403 ACCA 1406
 Db 1295 ACNA 1298

US-09-949-016-27
 ; Sequence 27, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 27
 ; LENGTH: 1715
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-27
 Query Match 40.0%; Score 656; DB 4; Length 1715;
 Best Local Similarity 71.9%; Pred. No. 1.2e-186;
 Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
 Qy 89 AGGAAGTCAAAATCTCGAACCAGAGGGTCCAGCCCATGTCCTCTGATTTCTTAGACA 148
 Db 367 AGGAGGCTGAAGCTGCTCGCTCCGCAACCAAGCCATATGTACCCCTCGGATTTCTGGATA 426
 Qy 149 AACTTATGGGGGCAACATCTGGATATGATGCGAGATTCGGCCCAATTTTAAAGGCCCA 208
 Db 427 AGCTAATGGGGGAGAACCTCCCGGATATGATGCGAGATTCAGGCCCAATTTTAAAGGTC 486
 Qy 209 CCGTGAAGCTGACCTGCAACATCTTCAACAAAGTTCAGCTCCGTCACCAAGACCAAC 268
 Db 487 CAGTGAACGCTGAGCTGCAACATTTTCAACAGCTTTGGTTCCATTTGCTGAGACAAC 546
 Qy 269 TGGACTACCGGTGAATGTCTTCTTGGGCAACAGTGGATGACCGACGCTGCTTAC 328
 Db 547 TGGACTATAGGTCAACATCTTCTCGGCGAGCAATGGAACGACCCCGCTGGCTATA 606
 Qy 329 GAGAATATCTGATGACTCTCTGGACCTCGATTCCTCATGCTGACTCTATCTGGAAGC 388
 Db 607 ATGAATACCTGACGACTCTCTGGACCTGGACCCATTCATGCTGACTCCATCTGGAAC 566
 Qy 389 CAGACTCTTCTTCTTCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCGGACAA 448
 Db 667 CTGACCTGTTCTTGGCAACGAGAGGGGGGCCACTTCCATGAGATCACCACAGACA 726
 Qy 449 AGTTACTGGCATCTTCAAGATGGGAATGTCTGTACAGCATCAGGCTGACCTCATTT 508
 Db 727 AATTGCTAAGGATCTCCCGGAATGGGAATGTCTCTACAGCATCAAGATCAACCTGAC 786
 Qy 509 TCTCTGCTGATGAGCTCAAGAACTTCCCATGAGATCCAGAGCTGCAAGATGACG 568
 Db 787 TGGCTTGGCCCATGAGCTTGAAGAAATTTCCCATGATGATGTCAGACATGATATGCAAC 846
 Qy 569 TTGAGAGCTCATCCATACTCTGAGCCCTCTGCCATCTCTGCTCACTTTTCAGTTGGCT 628
 Db 847 TGGAAAGCT-----TTGGATATA 864
 Qy 629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTCTGCTCAAGTGGCTGAGG 688
 Db 865 CGATGAATGACCTCATCTTGGATGGCAGGAACAGG---GAGCCGTGACAGGATGAGATG 921
 Qy 689 GGCTGACTGCTGCCAGTTTATCTTGGGATGAGAGGATCTAGGCTGTGTGACCAAGC 748
 Db 922 GACTAATCTGCCCCAGTTTATCTTGAAGGAAGAGAGGACTTGATGATGATGATGATG 981
 Qy 749 ACTACAACACAGGGAATTCACCTGATCGAGGTAAAGTTTCACTTGGAAACCGCAGATGG 808

982 ACTACAAACAGAGTAATTCACCTGATTTGAGGCCGGTTCACCTGGAGCGCAGATGG 1041
 809 GCTACTATCTGATTCAGATGTATATCCAGAGCTACTCATCGTCAATCTGCTGCTGCTGCT 868
 1042 GTTACTACCTGATTCAGATGTATATCCAGAGCTGCTCATTTGCTATCTCATGATCT 1101
 869 CTTCTGGATCAACATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
 1102 CTTCTGGATCAACATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
 929 TCACCATGACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
 1162 TCACCATGACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
 989 AGGCAATGACATCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
 1222 AAGCCATTGACATTTGGATGGCAGTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
 1049 ATGCTGCCATAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1108
 1282 ATGCTGCCATAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1341
 1109 AGAGGCGCAACGCTTGGAGGAAGATATCATCAAGAAAGTGTGCTTATTTCCGCTGCT 1168
 1342 GGAGCATCAAGAGGATGAAGCT-----GGAGAGCGCGCTTAACTTCTGCTGCT 1395
 1169 ATGGCTTGGGCC---ACTGCTGCGAGCAAGAGATGAGGTCCTCAATGGAAGGTTTC---TG 1222
 1396 ATGGGATGGCCAGCGCTGCTACAGGCCAAGATGCGATCTCAGTCAAGGGGGCCCAACA 1455
 1223 GATTTATAGTCCCACTCCAGCCCTCTTCTTAAGGGAAGAGAAACCAAGCGGAAC 1282
 1456 ACAGTAACACCAACCCCTCTGTCACCATCTAAGTCCCGAGGAGATGCGAAGAAC 1515
 1283 TCTAGTGACTGAGCAAGAGATTTGACACCATCTCCGGGCTGCTTCCCTTTCACCT 1342
 1516 TCTTATCCAGAGGCGCAAGAGATGCAAAATATCCCGATTTGCTTCCCGATGGCT 1575
 1343 TCTCATCTTCAATATCTTCTACTGCTGCTGCTTATTAAGTGTGCTGCTGCTGCTGCTGCTGCT 1402
 1576 TCTCATCTTCAATATCTTCTACTGCTGCTGCTTATTAAGTGTGCTGCTGCTGCTGCTGCTGCT 1635
 1403 ACCA 1406
 1636 ACAA 1639

RESULT 3

US-09-949-016-5267
 ; Sequence 5267, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: fastseq for Windows Version 4.0
 ; SEQ ID NO 5267
 ; LENGTH: 1854
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-5267

Query Match 39.5%; Score 648.6; DB 4; Length 1854;
 Best Local Similarity 70.0%; Pred. No. 2.2e-184;
 Matches 942; Conservativity 0; Mismatches 349; Indels 54; Gaps 3;
 63 CTTCTGCTGAGGCTGGCTTGGCAAAAGAGGAAGTCAAAATCTGGAACCAAGGGGTCACAGCC 122
 447 CTTTACAGGCGCTTCTGCAAAAGACCATGACTCCAGGCTCTGGAACCAACCTTTCACAGAC 506
 123 CATGTCCCTCTGATTTCTTAGACAAATTTATGGGGCGAAACATCTGGATATGATGCGCAG 182
 507 CCTATCTCTCTCAGATTTCTTGGACAAAGTTAATGGGAAGGACATCAGGATATGATCAAG 566
 183 GATTGGGCCCAATTTAAAGGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 567 AATCAGGCGCAAAATTTAAAGGCTCTCCAGTAAACGTTACTTGCATATTTTATCAACAG 626
 243 TTTTACGCTCCGTCACCAAGACCAATGAGTACCGGGTGAATGCTTCTTGGCGCAACA 302
 627 TTTTGGATCAGTCACAGAAACCGCATGACTACCGAGTGAATATTTTCTGAGACAACA 686
 303 GTGGAATGACCAACGCTCTCTACCGAATATTCCTGATGACTCTCTGAGACCTCGATCC 362
 687 GTGGAATGATTCACGCTGCGCTACAGTGAATACCCAGATGACTCCCTGACTTGGACCC 746
 363 CTCCATGCTGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTTAATGAGAAAGGGCCAA 422
 747 ATCCATGCTAGACTCCATTTGGAACCAAGTTGTTCTTGTGCAATTTTATCAACAG 806
 423 CTTTCAATGAGTGAACCAAGCAACAAAGTTACTTGCATCTTCAAGAAATGGGAATGTGCT 482
 807 CTTTCAAGATGTCACCACTGACCAAAATTTGCTACGGATTTTGGAAATTTGGAAGTGT 866
 483 GTACAGCATCAGGCTGACCTCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
 867 CTACAGTATCAGACTCACCTTGAACCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 543 GGACATCAGAGCTGACGATGAGGCTGAGAGCTCATCCATCTCTGAGCCCTCTGCTGCC 602
 927 GGATGCTCAGAGCTGTACAAATGAGCTGAGAGT----- 961
 603 ATCTCTGCTCAGTTTCAAGTGGCTTACACCATGAAAGACCTCGTGTGTTGAGTGGCTGGA 662
 962 -----TTGGGTACAGGATGATGACCTGATATTTGAGTGGTAAAGTGA 1004
 663 TGCTCTGCTGCTCAAGTGGCTGAGGGGCTGACTCTGCCCCAGATTTATCTTGGCGGATGA 722
 1005 TGGTC---CAGTGAAGTGTGCTGAAGATTGACCTGCCCCAGTTTATTTTGAAGAAGA 1061
 723 GAAGGATCTAGGCTGTGTACCAAGCACTACCAAGCAAGGAAATTCACCTGCTGCTGAGGT 782
 1062 GAAGGAACTTGGCTACTGTACAAAGCACTACCAAGCACTGGAAGAGTTTACCTGCAATGAGGT 1121
 783 AAAGTTTACCTGGAACGCGCAGATGGCTACTATCTGATTTCAGATGTACATCCCGACGCT 842
 1122 CAAATTTTCACTTGGAAAGCGCAATGGGATATTTATTTGATCCAGATGTACATCCCAAGCCT 1181
 843 ACTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
 1182 GCTTATAGTAAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
 903 TGTGGGCTTGGGCACTACCAAGCTGCTACCATGACCAAGCAAGCTGCTGCTGCTGCTGCTGCT 962
 1242 GGTGCACTTGGGCACTACCAAGCTGCTTAAAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
 963 CTCCTTGGCTTAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 1302 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
 1023 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
 1362 GTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421
 1083 AGAATTCATCAGCTTTCGAAGAGGAGGAGGCGCCCAACGCTTGGAGGAAGATATCATCCA 1142

Db 1422 GGAGTTCCTGCGCCTCCGGAAGACAGAGAGGCGAATAAGGAAGAAGACGTTACTCG 1481
 Qy 1143 AGAAAGTCGTTCTATTTCGTTGGCTATGCTTGGCCACTGCTGCGCAGGAGAGATCG 1202
 Db 1482 TGAAGTCGTTTAAATTTAGCGGTTATGGGATGGGTCACTGCTCCAAAGTGAAGATCG 1541
 Qy 1203 AGTCCAAATGGAAGGTTCTGGCATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGA 1262
 Db 1542 ACAGCTGTCAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCAAGAA 1592
 Qy 1263 AGAGAAACCCAGCGGAACCTACGTGAGATGAGCAAGAGAAATTCACACCATCTCCG 1322
 Db 1593 TGGAGATGCTATCAAGAAGAAGTTTGTGACCGGGCAAAAGGATTGACACGATATCTCG 1652
 Qy 1323 GCGTGTCTCCCTTTCACATCTTCTCATCTTCATATCTTCTACTGGGTTGTCTATAAGT 1382
 Db 1653 AGCTGCTTCCCATGGCCCTTCTCATTTTCAACATCTTTTACTGGATCACATACAAGAT 1712
 Qy 1383 GCTATGGTCAGAAGATATCCACCAG 1407
 Db 1713 CATTCGGCATGAGATGTCACAAG 1737

RESULT 4

US-09-016-434-271
 ; Sequence 271, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 271:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 253 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNOT23
 ; CLONE: 1695770
 ; US-09-016-434-271

Query Match 15.4%; Score 253; DB 4; Length 253;

Best Local Similarity 100.0%; Pred. No. 7.9e-66;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1004 GGATGCTGTGTCTGCTCTTTGTTGTTGCTTGGCTTGTGCTGAGTATGCTGCCATAAAATT 1063
 Db 1 GGATGCTGTGTGTCTGCTCTTTGTTGTTGCTTGGCTTGTGCTGAGTATGCTGCCATAAAATT 60
 Qy 1064 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAAGGAGAGCGCCAACGCT 1123
 Db 61 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAAGGAGAGCGCCAACGCT 120
 Qy 1124 TGGAGAAAGATATCATCAAGAAAGTCGTTTCTATTTCCGTCGCTATGCTTTGGGCCACT 1183
 Db 121 TGGAGAAAGATATCATCAAGAAAGTCGTTTCTATTTCCGTCGCTATGCTTTGGGCCACT 180
 Qy 1184 GCTCGAGCAAGAGATGAGGTCCAAATGGAAGGTTCTGGCATTTATAGTCCCCACCTC 1243
 Db 181 GCTCGAGCAAGAGATGAGGTCCAAATGGAAGGTTCTGGCATTTATAGTCCCCACCTC 240
 Qy 1244 CAGCCCTCTTCT 1256
 Db 241 CAGCCCTCTTCT 253

RESULT 5

US-09-949-016-3698
 ; Sequence 3698, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3698
 ; LENGTH: 2109
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-3698

Query Match 12.9%; Score 211; DB 4; Length 2109;
 Best Local Similarity 55.8%; Pred. No. 1.3e-52;
 Matches 498; Conservative 0; Mismatches 345; Indels 50; Gaps 3;
 Qy 170 GATATGATCCAGGATTCGGCCCAATTTTAAAGGCCACCCGTCGAGACGTGACCTGCAACA 229
 Db 273 GTTATGATCCAGGATAAGACCAACTTCAAAGGCATTCTGTGTAGTAGTAGTCAACA 332
 Qy 230 TCTTCATCAACAGTTTCAGTCCGTCGTCACCAAGACCAAT--GGACTACGGGTGATGT 287
 Db 333 TTTTATTAAACAGTTTGGATCCATTCAATATCTTTTCTTCAGGACTATAGAGTTAAACAT 392
 Qy 288 CTTCTTGGCGCAACAGTGGAAATGACCCAGCCTGCTCTACCGAGAAATATCTTGAT----- 342
 Db 393 CTTCTGAGACAAAATGGAATGACCCAGGCTGAAGCTCCCGAGTATTTAGGGGTTTC 452
 Qy 343 -CACTCTCTGGACCTCGATCCCTCCATGCTGAGACTCTATCTGGAAGCCAGACCTCTTCTT 401
 Db 453 AGATGCACTGACAGTGGATCCAAATGTACAAGTGTATGGAACCTGATTTATTTT 512
 Qy 402 TCGTAATGAGAAAGGGGCCAACTTCCATGAGTGCACCGACCAAGTACTTACTGCGCAT 461
 Db 513 TCGAAATGAAAAAAGTGCCCAATTTTTCATGATGTGACCCAGGAAACATCTCTCTCTTAT 572

```

QY 462 CTTCAAGATGGGAATGCTGTACAGCATCAGCTGACCTCATTTTGTCTCCTGCTGAT 521
Db 573 TTTTGTGTGGAGATGCTCTTGTACAGCATGAGGTATCTATTTACCTTTTCATGCCCCCTT 632
QY 522 GGACCTCAAGAACTTCCCATGACATCCAGACGTCACGATCGAGCTTGAGAGCTCATC 581
Db 633 GGACTTGACATTTTCCCATGATACACACGCTTGCAGATGCAACTGGAGACT---- 688
QY 582 CATACTCTGACGCTCTGCTGCTCTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Db 689 -----TTGGTTACACAACTGATGATTT 710
QY 642 CGTGTGTTGAGTGGAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
Db 711 ACATTTATCTGCGAGTCAGAGATCTGCTGCAATAGAAAATTTGCTTGTGCTCAAT 770
QY 702 CCAGTTTATCTTCCGGGATGAGAGATCTAGGCTGTGTACCAAGCACTTACACACAGG 761
Db 771 TGATATCAAAAGAGATATTGAATATGTTAATCTGTACAAATATCTATAAGGCAAGG 830
QY 762 GAAATTCACCTGATCGAGTAAGTTTCACTGGAACGCGAGATGGGCTTACTATCTGAT 821
Db 831 CTACTACACATGCTGGAAGTCTATCTTCACTGAGGAGGAGGCTGGCTTTTACATGAT 890
QY 822 TCAGATGTATATCCCGAGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881
Db 891 GGGGGTCTAGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950
QY 882 CATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941
Db 951 CCGGAGCGAGTGTGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
QY 942 CCAGAGCTTGGCTCCCGGCTCTTTTGGCTAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 1011 TGAGTGACACACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
QY 1002 CTGGATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
Db 1071 TTGGCTTATGCTTGGCTTCTTCTTGGGTTTGTCTTCCCTGGTGGAGTATGCG 1123

```

RESULT 6

```

US-08-809-802-11
; Sequence 11, Application US/0809802
; Patent No. 6455276
; GENERAL INFORMATION:
; APPLICANT: Le Bourdelles, Beatrice
; APPLICANT: Whiting, Paul John
; TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT
; TITLE OF INVENTION: OF THE GABA-A RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,802
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02323
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Yablonsky, Michael D
; REGISTRATION NUMBER: 40,407

```

```

; REFERENCE/DOCKET NUMBER: T1292
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-4678
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1555 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 47...1402
; OTHER INFORMATION:
; US-08-809-802-11

Query Match 11.8%; Score 193; DB 3; Length 1555;
Best Local Similarity 53.3%; Pred. No. 2.9e-47;
Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

QY 142 CTAGACAAACTTATGGGGGAAACATCTGGATATGATCCAGGATTCGGCCCAATTTTAA 201
Db 167 CTCCTCCAACTGGACGGGCTGTAGCGGTACGCCGCAACTTCGGGCTGGCATCGGA 226
QY 202 GGGCCCAACCGTGAACTGACCTGCAACATCTTTCATCAACAGTTTTCAGTCCGTCACCAAG 261
Db 227 GGGCCCCCGTGAATGTGGCCCTTGGCCCTGGAGGTGGCCAGCATCGACACATCTCAGAG 286
QY 262 ACCACAAATGGACTACCGGGTGAATGTCTTCTTGGGCAACAGTGGGAATGACCCACGCTG 321
Db 287 GCCAAATGGAGTACACCATGACGGTGTCTTCTGACCAAGAGCTGGCGGAGCAGCAGGCTC 346
QY 322 TCTACTCCGAGAAATATCTCTGATGACTCTCTGGACCTCGATCCCTCTCATCTGGACTCTATC 381
Db 347 TCTACT---AACCAACAAACGAGACCTTGGGCTGGACAGCGCTTCGTGGACAAGCTG 403
QY 382 TGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCCAACTTCATCAGAGTGACCAAG 441
Db 404 TGGCTGCCGACACCTTTCATCGTGAACGCAAGTGGGCTGGTTCCACGACGTGACGGTG 463
QY 442 GACACAAAGTTACTGGCGATCTTCAAGAAATGGGAATGTCTGTACAGCATCAGGCTGACC 501
Db 464 GAGAACAAAGCTCATCGGCTGACGCCGCGAGGGGTGATCTCTGTACAGCATCCGAATCACC 523
QY 502 CTCATTTTGTCTCTGCTGATGACCTCAAGAACTTCCCATGGACATCCAGAGCTGCACG 561
Db 524 TCACATGTGGCTGCGACATGGACCTGGCCAAATTCCTCATGGACGAGCAGGAGTGCATG 583
QY 562 ATGACAGCTTGAGAGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTTCAGTT 621
Db 584 CTGGACCTGGAGACT-----AC 601
QY 622 GGCTACACCATGAAGACCTCTGTGTGAGTGGCTGGAAGATGCTCCTGTCTGCTCAAGTG 681
Db 602 GGTACTCATCGAGGACATCGTCTACTACTTGTGCGAGAGCCAGGAGCAGCATCCACGGG 661
QY 682 GCTGAGGGCTGACTCTGCCCGAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTGT 741
Db 662 CTGGACAAGCTGAGCTGGCGAGTTTCCATCATCCAGCTACCGCTTACACAGGAGCTG 721
QY 742 ACCAAGCACTACAAACAGGGAAATTCACCTGTGATCGAGGTAAGTTTTCACCTGGAAACGG 801
Db 722 ATGAACCTCAAGTCCGCTGCCAGTTTCCCAAGCTCAGCTGACATTCACCTGCGGAGG 781
QY 802 CAGATGGGCTACTATCTGATTCAGATGTACATCCCGAGCCTACTCATCTGTCATCTGTC 861
Db 782 AACCGCGGGTGTACATCATCTCAATCTACATGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 841
QY 862 TGGGTCTCTTCTGGATCAACATGGATGCTGCCCTGCGGCTGGGCTGGGCTGGGCTGCTGCTGCT 921
Db 842 TGGGTCTCTTCTGGATCAGCCAGGCGGGGTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGCT 901

```

```

Qy 922 ACCGTGCTCACCATGACCAACCGAGCTCTGGCTCCGGGGCCTCTTTGCCCTAAGGTGTCC 981
Db 902 ACGGTGCTGACGATGACCAACCGAGCTCTGGCTCCGGGGCCTCTTTGCCCTAAGGTGTCC 961
Qy 982 TACGTGAAGGCAATCGACATCTGGATGGCTGTGTGTCTGTCTTTTGTGTTGGCTGCCTTG 1041
Db 962 GCCATCAAGGCACCTGGAGCTTACTTCTGGATCTGCTATGTCTTCGTGTTTGGCCGCTG 1021
Qy 1042 CTGGAGTATGC 1052
Db 1022 GTGGAGTACG 1032

RESULT 7
US-09-592-891A-13
; Sequence 13, Application US/09592891A
; Patent No. 6329174
; GENERAL INFORMATION:
; APPLICANT: Xiao-Zhou Michelle Wang
; APPLICANT: Xavier Georges Sarda
; APPLICANT: Michael David Tomalski
; APPLICANT: Vincent Paul Mary Wingate
; TITLE OF INVENTION: Heliothis Glutamate Receptor
; FILE REFERENCE: A32815 072667.0118
; CURRENT APPLICATION NUMBER: US/09/592,891A
; CURRENT FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 4621
; TYPE: DNA
; ORGANISM: Heliothis virescens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (764)...(764)
; OTHER INFORMATION: n is a, t, g, or c
; US-09-592-891A-13

```

Query Match	11.4%	Score 187;	DB 3;	Length 4621;
Best Local Similarity	54.5%;	Pred. No. 3.6e-45;		
Matches	472;	Conservative 0;	Mismatches 346;	Indels 48; Gaps 3;
Qy	208	CCCGTGAACGTGACCTCACAATCTTCATCAACAGTTTCAGTCCGTCACCAAGACCACA	267	
Db	330	CCACGGTAGTACGGCTCAATATATTTGTCGGAAGTATATCAAAGATCGATGAGTCA	389	
Qy	268	ATGACTACCGGGTGAATGTCTTCTTGGCGCAACAGTGGAAATGACCCACGCTCTCCATC	327	
Db	390	ATGGAATACTCCGTACAGTTAAACGTTTCGGGAACAATGGTTAGATGAACGGCTCAAATTC	449	
Qy	328	CGAATAATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGAGTCTATCTGGAAG	387	
Db	450	AATAATCTTGGAGTGCCTCAAAATACCTGCACACTGACCGAAGCAACAGAGTCTGGATG	509	
Qy	388	CCAGACCTCTTTTGTGTTAATGAAAGGGGCCAACTTCCATGAGGTGACCAACGGAAC	447	
Db	510	CGTATCTATTTCTCTCAACGAGAAGAGGTCAITTTCCACAACATCATGATCCGAAC	569	
Qy	448	AAGTTACTGGGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATCAGGCTGACCCCTCAT	507	
Db	570	GTGTACATCCGATCTTCCCAACGGCAACGTGCTGTACAGCATCCGAATCTCCTTGACG	629	
Qy	508	TTGTCTCCCTGATGGACCTCAAGAACTTCCCAATGGACATCGACGTCGACGATCGAG	567	
Db	630	CTCTCGTGGCCCATGAACCTCAAGTTGTGATCCCTCGGATAAGCAGACCTGTCTCGCTCAGG	689	
Qy	568	CTTCGAGAGCTCATCCATACCTGCGAGCCCTCTGCCACTCTCTGTACATTTTCAGTTGGCTAC	627	
Db	690	ATGGCTAGTT-----ATGGTTGG	707	
Qy	628	ACCATGAAGAACCCTCGTGTGTTTGGTGGCTGGAAGATGCTCCTGCTGTCCAAAGTGGCTGAG	687	

Db	708	ACACAGACACTAGTGTTCCTATGGAAGGAGCGACC--CGGTGCAGGTGGTGAAN	764
Qy	688	GGGCTGACTCTGCCCCAGTTTATCTTTGCGGGATGAGAAGGATCTAGGCTGTTTGTACCAAG	747
Db	765	AACTTACACCTGCTCGGTTTCAC--GCTGGAGAAGTTCCTCACTGACTACTGCAACAGT	821
Qy	748	CACATACACACAGGAAATTCACCTGATCGAGGTAAAGTTTCACTGGACCGCAGATG	807
Db	822	AAGACTAATACCGGTGAATACAGTTGCTGGAAGGTAGACTTGTCTTTCAAACGCGAGTTC	881
Qy	808	GGCTACTATCTGATTCAGATGTACATCCCGAGCTACTCATCTCTCATCTGTCTCTCGGTC	867
Db	882	AGTTACTACTGATTCAGATCTACATTCGCTGCTGCATGCTGTGTCATCGTGTCTCGGGTG	941
Qy	868	TCCTTCTGGATCAACATGGATGCTGCGCCCTGCGCCGTGTGGCCCTGGGCACTACACACCGTG	927
Db	942	TCCTTCTGGCTGGACGAGGAGCTGTGCCTGCGAGGCTCTCACTAGGAGTGACGACTTTA	1001
Qy	928	CTCACCATGACACCCAGAGCTCTGGCTCTCCGGGCTCTTTTGGCTAAGGTGTCCTACGTC	987
Db	1002	CTTACAATGGCAGCCAGTCGTGAGGATCAACCGCTCCCTACACCGGTGTCTCTACACG	1061
Qy	988	AAGCAATCACATCTGGATGCTGTGCTCTCTTGTGTGTCTGCTGCTGCTGCTGCTGGAG	1047
Db	1062	AAAGCCATTCAGCTGGAATGCTGTATGTCACATTCGTAFTCGGAGCGCTACTAGAG	1121
Qy	1048	TATGCTGCCATAAATTTGTTTCTCG	1073
Db	1122	TTGCGGCTCGTCAACTATGCGTCTCG	1147
RESULT 8			
US-09-969-844-13			
; Sequence 13, Application US/09969844			
; Patent No. 6780601			
; GENERAL INFORMATION:			
; APPLICANT: Xiao-Zhou Michelle Wang			
; APPLICANT: Xavier Georges Sarda			
; APPLICANT: Michael David Tomalski			
; APPLICANT: Vincent Paul Mary Wingate			
; TITLE OF INVENTION: Heliothis Glucamate Receptor			
; FILE REFERENCE: A32815-1 072667.0178			
; CURRENT APPLICATION NUMBER: US/09/969,844			
; CURRENT FILING DATE: 2001-10-03			
; NUMBER OF SEQ ID NOS: 14			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 13			
; LENGTH: 4621			
; TYPE: DNA			
; ORGANISM: Heliothis virescens			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (764)...(764)			
; OTHER INFORMATION: n is a, t, g, or c			
US-09-969-844-13			
Query Match			
Best Local Similarity 11.4%; Score 187; DB 4; Length 4621;			
Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;			
Qy	208	CCCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTTCAGCTCCGCTCACCAGACACACA	267
Db	330	CCAGCGGTAGTCGCTCAATATATTGTCGGAAGTATATCAAGATCGATGACGTCACA	389
Qy	268	ATGAGCTACCGGTGAATGTCCTTTGCGGCAACAGTGAATGACCCAGCCCTGTCTCTAC	327
Db	390	ATGGAATACTCCGTACAGTTTAAAGTTTCGGGAACAATGGTTAGATGAACGGCTCAAATTC	449
Qy	328	CGAGATATCTGTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATCTGGAAG	387
Db	450	AATAATCTTGGAGGTGCGCTCAAAATACCTTGACACTGACCGAAGCAACAGAGTCTGGATG	509
Qy	388	CCAGACTCTCTTTTGTGAATGAGAAGGGGGCCAACTTCATGAGGTGACCAACCGGACAAC	447


```
Db 510 CCTGATCTATTCTTCCACGAGGAGGTCATTTCCACAAATCATCATGCCGAAC 569
Qy 448 AGATTACTGGCATCTTCAAGATGGAATGTGCTGTACAGCATCAGGCTGACCCCTCAT 507
Db 570 GTGTACATCGGATCTTCCCAACGCAAGTGTGTACAGCATCGAATCTCCTTGACG 629
Qy 508 TTGTCTGCTGTATGATGATCTTCCCAATGGAATCTCCCATGAGCATCAGAGTGCAGATGAG 567
Db 630 CTCCTGCTGCCCCAAGCACTCAAGTTGTATCCCCCTGGTAAGCAGACCTGCTGCTCAGG 689
Qy 568 CTTGAGAGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTTCAGTTGGCTAC 627
Db 690 ATGGCTAGTT-----ATGGTTGG 707
Qy 628 ACCATGAAGACCTCTGTTGAGTGGCTGGAAGATGCTCTGCTGTCCAAGTGGCTGAG 687
Db 708 ACCACAGACGACTTAGTGTCTTATGGAAGGAGCGACC---CGGTGCAGGTGGTGAAN 764
Qy 688 GGGCTGACTCTGCCCGCAGTTATCTTGGCGGATGAGAAGATCTAGGCTGTTGTACCAAG 747
Db 765 AACTTACACTCTGCTCGGTTCAC---GCTGGAGAAGTTCTCTACTGACTACTGCAACAGT 821
Qy 748 CACTCAACACAGAGGAAATTCACCTGCATCGAGGTAAAGTTTCCACCTGGAACGCGAGATG 807
Db 822 AGACTTAATACCGTGATACAGTTGCTGGAAGTGAATCTGCTCTTCAACCGAGTTTC 881
Qy 808 GGCTACTATCTGATTCAGATGTACATCCCGAGCCTACTCATGTCATCTCTGCTGCTGGTC 867
Db 882 AGTTACTACTGATCCAGATCTATATTCGTGCTGTCATGCTGTCATGCTGCTGGTG 941
Qy 868 TCCTTCTGATCAACATGATGCTGCCCTGCCCGTGGCGGCTGGGATCACCACCGTG 927
Db 942 TCCTTCTGGCTGACGAGAGTGTGCTGCGAGGGTCTCACTAGGAGTGACGACTTTA 1001
Qy 928 CTCACCATGACCAACAGAGCTCTGCTCCCGGCTCTTTGCTAAGGTGCTACGTCG 987
Db 1002 CTTACAATGGCAGCCAGTGTGAGGATCAACCGTCCCTACCAACCGGTGCTACAGC 1061
Qy 988 AAGGCAATCGACATCTGGATGGTGTGTCTGCTCTTTGTTGCTGCTGCTGCTGCTGAG 1047
Db 1062 AAAGCATTGAGCTGTGGACTGGTGTATGTCTCACATTCGTATTCGAGGCTACTAGAG 1121
Qy 1048 TATGTCGCATAAATTTGTTCTCG 1073
Db 1122 TTGCGCTGCTCAACTATGCTGCTCG 1147
```

RESULT 9

US-08-417-330A-15
; Sequence 15, Application US/08417330A
; Patent No. 5719057

GENERAL INFORMATION:

; APPLICANT: HADINGHAM, KAREN
; APPLICANT: LE BOURDELLES, BEATRICE
; APPLICANT: WHITING, PAUL
; APPLICANT: WINGROVE, PETER
; TITLE OF INVENTION: STABLY TRANSFECTED CELL LINE EXPRESSING
; TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLONED
; TITLE OF INVENTION: GABA-RECEPTOR SUBUNIT CDNA SEQUENCES
; NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

; ADDRESSEE: J. MARK HAND - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ

COUNTRY: US

; ZIP: 07065-0900

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq For Windows Version 2.0

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,330A
; FILING DATE: 05-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HAND, MARK
; REGISTRATION NUMBER: 36,545
; REFERENCE/DOCKET NUMBER: T11091A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; TELEFAX: 908-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 298...1683
; OTHER INFORMATION:
; US-08-417-330A-15

Query Match 11.1%; Score 182.2; DB 1; Length 2310;
Best Local Similarity 54.0%; Pred. No. 6.5e-44;
Matches 488; Conservative 0; Mismatches 373; Indels 42; Gaps 4;
Qy 168 TGGATATGATGCCAGGATTCGGCCCAATTTTAAAGCCCAACCGTGAACCTGACCTGCAA 227
Db 471 TGGCTACGACAAACAGACTTCGGCCCGGCTGGGAGAGCGCATCCTCAGCTGAGGACCGA 530
Qy 228 CATCTTCATCAACAGTTTCAGCTCCGTACCAAGACCACAAATGGACTACGGGTGATGT 287
Db 531 CATCTACGTACCAACAGCTTCGGCCCGGTGTCGACACGSAATGGAGTACACCATGACGT 590
Qy 288 CTTCTTGGCGCAACAGTGGAAATGACCCACGCTGTCTACCGAGAAATATCTCTGATGACTC 347
Db 591 GTTTTTCGACAAAGCTGGAAAGATGAAAGGCTTCGGTTTAAGGGG---CCCATGCAGCG 647
Qy 348 TCTGGACCTCGATCCCTCCATCTGGAATCTATCTGGAAGCCAGACCTCTTTTGTGCTAA 407
Db 648 CCTCCCTCTCAACACCTCTTGGCCAGCAAGATCTGGACCCCAACAGCTCTCGGCTGGAGGA 707
Qy 408 TGAGAAAGGGGCAACTTCCATGAGGTGACCAAGCAACAAAGTTACTTGGCATCTTCAA 467
Db 708 CGGGAAGAAAGTCCATCGCTCACAACATGACCAACGCGCCCAACAAAGCTCTCGGCTGGAGGA 767
Qy 468 GAATGGGAATGTGTGTACAGATCAGGCTGACCCCTCATTTTGTCTGCTGCTGATGACCT 527
Db 768 CGACGCAACCTCTCTACACCATGCGCTTGACCATCTCTGCAGAGTGCCCATGACGT 827
Qy 528 CAAGAACTTCCCATGAGCATCCAGACGTGACGATGACAGCTTGAGAGCTTCATCCATCT 587
Db 828 TGAGGACTTCCCGATGATGCGCAGCTTGCCTCTGAAATTTGGCAGCT-----877
Qy 588 CTGACGCCCTCTGCCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAGACCTCTGTT 647
Db 878 -----ATGCGTACCCCTAAATCTCTGAAGTCTGTTTACGCTGAGCAACACGCTCCAC 926
Qy 648 TGAGTGGCTGGAAGATGCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTT 707
Db 927 CAAGTCGTGGTGGTGGC-----GGAAGATGGCTCCAGTCCAGTGAACCACTA 971
Qy 708 TATCTTGGCGGATGAGAAGGATCTAGGCTGTTGTACCAAGACACTACACACAGGGAAAT 767
Db 972 CCACCTGATGGGCGAGACG---TGGGCACTGAGACATCAGCACCCAGCAGCAGCGCAATA 1028
Qy 768 CACCTGCATCGAGGTAAAGTTTTCACCTGGAAACGGGAGATGCTTCTATCTGATTCAGAT 827


```
Db 394 ATGGAATACCTCGTACAAATTAACGTTTCGGGAACAATGTTAGATGAACGGCTCAAAATTC 453
Qy 328 CGAGAATACTCTGATGACTCTCTGGAACCTCGATCCCTCCATGCTGAGCTCTATCTGGAAG 387
Db 454 AATAAATCTTGGAGGTGCGCTCAAAATACCTGACACTGACTGAAGCCAAAGAGTCTGGATG 513
Qy 388 CCAGAGCTCTCTTTCTGTAATGAGAAGGGGCCAACTTCCATGAGGTGACCAAGGACAC 447
Db 514 CTTGATCTATTCTTCTCCAAAGGAAGGAGTCAATTCACAAACATCATCATGCCGAAC 573
Qy 448 AAGTTACTGCGCATCTTCAAGAAATGGAATGTCTGTACAGCATCAGGCTGACCCCTCAIT 507
Db 574 GTGTACATCGAATCTTCCCAACGCAAGTGTGTACAGCATCGAATCTCCCTGACG 633
Qy 508 TTGCTCTGCTGATGAGCACTCAAGAAATCTCCCATGAGCATCCAGACGTGCGAGTGCAG 567
Db 634 CTCTCGTCCCAATGAACCTCAAGTTGTACCCCTCGAATGAAGCAGACCTGCTCGCTCAGG 693
Qy 568 CTTGAGAGCTCATCCATACCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTTGGCTAC 627
Db 694 ATGGCTAGTT-----ATGGTTGG 711
Qy 628 ACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGCTGTCCAAAGTGGCTGAG 687
Db 712 ACCACAGACGACTTAGTGTCTTATGGAAGGAAGGCGACG---CGGTGAGGTGGTGA 768
Qy 688 GGCTGACTCTGCCCCAGTTTATCTTTCGGGATGAGAAGGATCTAGGCTGTTGTACCAAG 747
Db 769 AACTTACACTGCTCGGTTTCAC--GCTGGAGAAGTTCCTCACTGACTACTGCAACAGT 825
Qy 748 CACTACACACAGGGAATTCACCTGATGAGGTAAAGTTTCACTGGAACGGGAGATG 807
Db 826 AAGACTAATACCGGTGAATACAGTTGCTGGAAGTAGACCTGCTCTTCAAACCGGAGTTC 885
Qy 808 GGCTACTATCTGATTCAGATGTACATCCCGAGCTACTCATCTCATCTGCTCGTGGTC 867
Db 886 AGTTACTACTGATTCAGATCTACATTCGCTGCTGATGCTGATGCTGATGCTGCTGGGTG 945
Qy 868 TCCTTCTGATCAACATGATGATGCTGCCCTGCTGCGGCTGCGGCTGCGGCTACCAACCGTG 927
Db 946 TCCTTCTGCTGGAACCGAGGCTGTGCTGCGAGGCTCTCACTAGGAGTGAAGTCTTA 1005
Qy 928 CTCACCATGACACCCAGAGCTCTGGCTCCCGGGCTCTTTGCTGAAGGTGCTCAAGTG 987
Db 1006 CTTACAATGGCGACCCAGTCTGAGGATCAACGCTGCCCTACCAACCGGTGCTCTACAG 1065
Qy 988 AAGGCAATCGACATCTGGATGGCTGTGTCTGCTCTTTGTTGCTGCTGCTGCTGCTGAG 1047
Db 1066 AAGCCATTGATGCTGGAGTGGGTATGTCTCACATTCGATTCGGAGCGCTACTAGAG 1125
Qy 1048 TATGCTGCCATMAATTTTGTCTCG 1073
Db 1126 TTGCGCTCGTCAACTATGGTCTCG 1151
```

RESULT 12

```
US-08-435-933-5
; Sequence 5, Application US/08435933
; Patent No. 5693492
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Arena, Joseph P.
; APPLICANT: Paresse, Philip S.
; APPLICANT: Liu, Ken K.
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
; TITLE OF INVENTION: CHANNELS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
```

```
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,933
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wallen, III John W.
REGISTRATION NUMBER: 35,403
REFERENCE/DOCKET NUMBER: 19264
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3905
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-435-933-5
```

```
Query Match 10.9%; Score 179; DB 1; Length 3958;
Best Local Similarity 53.7%; Pred. No. 8.4e-43;
Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;

Qy 193 AATTTTAAAGGCCACCCGCTGAACTGACCTGCAACATCTTCAATCAACAGTTTCAGCTCC 252
Db 570 AATGGCACAGATGGTCCGCCATAGTCAGAACTATCTATCTGTCGCGATATTATGACG 629
Qy 253 GTCACCAACACACATGCACTACCGGGTGAATGCTCTTCTGCGGCAACAGTGAATGAC 312
Db 630 ATTAGTGATTTAAATGAGGATACAGTGTGAGTTAACTTCCTGCGTGAACAGTGGAGCGAT 689
Qy 313 CCACGCTCTCTTACCGAATATCTGATGACTCTCTGGAACCTCGATCCCTCCATGCTG 372
Db 690 GAACGCTCAAGTTCGACGATATCCAGGCTCGCTAAAGTATCTGACCTGACGGAGCG 749
Qy 373 GACTCTATCTGAAGCCAGACCTCTTTCTGTAATGAGAAAGGGGCCAACTTTCATGAG 432
Db 750 AACCAGGTGTGATGCGCGATCTTTTCTCTCGAACGAGAGGAGGACACTTCCCAAC 809
Qy 433 GTGACACCGGACCAACAAAGTTACTGCGCATCTTCAAGAATGGGAATGCTGTGACAGATC 492
Db 810 ATCATCATGCCCAATGTGTATATTCGCATCTTCCCAACGGATCTGTGCTATATAGTATA 869
Qy 493 AGGCTGACCCCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAG 552
Db 870 CGTATCTCGCTGACATTTGGCTGCCCTGCCAATGAACCTGATATCGCTGGATAGACAG 929
Qy 553 ACGTGCAAGATGAGCTTGAAGAGCTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCA 612
Db 930 ATCTGCTACTACGATGGCCAGCT-----954
Qy 613 CTTTCAGTTGGCTACACCATGAAGACCTCTGTTTGTGCTGGCTGGAAGATGCTCTGCT 672
Db 955 -----ATGGCTGGACCAACCAAGCTTGGTCTTCTCTGGAAGAGGAGGCGATC---CC 1004
Qy 673 GTCCAAGTGGCTGAGGGGCTGACTCTGCGCCCAAGTTTATCTTGGGATGAGAAGGATCTA 732
Db 1005 GTACAGTGGTAAAGNACTTACACCTACCTCGCTTCACTGGAGAGTTTCTGACT--- 1061
Qy 733 GGCTGTTGACCAAGCACTACCAACAGAGGAATTCACCTGCAATCAGAGTAAAGTTTCA 792
Db 1062 GATTACTGTAAACAGTAAACCAACACCGGTGAATACAGTTGCTCAAGTGCATCTACTA 1121
Qy 793 CTGGAAACGCGACATGGGCTACTATCTGATTCAGATGTATATCCCAAGCTTACTCATCTC 852
```

Db 1122 TTCAGCGAGAAATCTCATATTACTTAATAACAAATTTATATACCATGCTGTATGTTGGTC 1181
 QY 853 ATCTGTCTGGGTCTCTCTTCTGGATCAACATGATGCTGCCCTCCCGTGTGGCGCTG 912
 Db 1182 ATTGTATCATGGTATCATCTCTGGTGGATCAAGAGAGATGACCGGCGGAGTGTCACTG 1241
 QY 913 GGATCACCACCGTGTCTACCATGATGACCCAGAGCTCTGGTCCCGGCGCTCTTTGGCT 972
 Db 1242 GGTGTCAACACCGTGTCTGACCATGCGCCAGAGCGTGGGCATAAACGCTCCCTGCCG 1301
 QY 973 AAGTGTCTCATGTAAGCAATCGACATCTGGATGGCTGTGTCTCTTTGTGTTTC 1032
 Db 1302 CCCGTTTCTATACGAAGCCATCGATGTGTGACAGCGGTGTGTCTGACGTTCTGTTTC 1361
 QY 1033 GCTGCTTCTGAGTATGCTGCCATAAATTTTGTTCG 1073
 Db 1362 GGGGCCCTGCTGAGTTCGCCCTGGTGAATGATCCG 1402

RESULT 13

PCT-US96-06035-5
 ; Sequence 5: Application PC/TUS9606035
 ; GENERAL INFORMATION:
 ; APPLICANT: Cully, Doris F.
 ; APPLICANT: Arena, Joseph P.
 ; APPLICANT: Pares, Philip S.
 ; APPLICANT: Liu, Ken K.
 ; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
 ; TITLE OF INVENTION: CHANNELS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Jody M. Giesser
 ; STREET: 126 East Lincoln Avenue - P.O. BOX 2000-0907
 ; CITY: Rahway
 ; STATE: New Jersey
 ; COUNTRY: US
 ; ZIP: 07065-0907
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/06035
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Giesser, Jody M.
 ; REGISTRATION NUMBER: 32,838
 ; REFERENCE/DOCKET NUMBER: 19264 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (908) 594-3046
 ; TELEFAX: (908) 594-4720
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3958 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; PCT-US96-06035-5

Query Match 10.9%; Score 179; DB 5; Length 3958;
 Best Local Similarity 53.7%; Pred. No. 8.4e-43;
 Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;
 QY 193 AATTTTAAAGGCCACCCGTGACGTGACATCTTCATCAACAGTTTTCAGCTCC 252
 Db 570 AATGGCACAGATGTTCCCGCCATGATCAATCTATTCTGTCGAGTATTATGACG 629
 QY 253 GTCACCAAGCCAAATGACTACCGGGTGAATGTTCTTTCGGGCAACAGTGGAAATGAC 312
 Db 630 ATTAGTGATTAATAAATGGAGTACAGTGTGCAGTTAACTTCGTTGAACAGTGGACGGAT 689

QY 313 CCACGCTGTCTTACCGAGAAATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTG 372
 Db 690 GAACGCTCAAGTTCGACGATATCCAGGTCCCTAAAGATATCTGACCTCGAGGAGCG 749
 QY 373 GACTCTATCTGAAGCCAGACCTCTTTCTTGTAAAGAAAGGGGCCAACTTCCATGAG 432
 Db 750 AACCCGCTGTGATGCCGATCTTTTCTCTGAAAGAGAGGAGGACACTTCCACAAC 809
 QY 433 GTGACCAAGCAACAAAGTTACTGCGCATCTTCAAGAAATGGAATGTGCTGTACGATC 492
 Db 810 ATCATCATGCCCAATGTGTATATTCGCATCTTCCCAACGGAATCTGTGTATATAGTATA 869
 QY 493 AGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCCATGACATCTGTA 552
 Db 870 CGTATCTCGTCACATTTGGCTGCCCAATGAACCTTAAGCTGATCCGCTGGATAGACAG 929
 QY 553 ACGTCAGATGCGCTTGAGAGCTCATCATCTCTGAGCCCTCTGCCATCTCTGTCTCA 612
 Db 930 ATCTGCTCACTACGATGGCCAGCT----- 954
 QY 613 CTTTCAGTTGGCTACACCATGAAGACCTCTGTTTGTAGTGGCTGGAAGATGCTCTGCT 672
 Db 955 -----ATGGCTGGACCAACCAAGCTTGGTCTTCTGTGGAAGAGGCGGATC---CC 1004
 QY 673 GTCCAAGTGGCTGAGGGGCTGACTCTGCCCAAGTTTATCTTCCGGGATGAGAAGGATCTA 732
 Db 1005 GTACAGTGGTAAAGAACTTACACCTACTCTGCTTCACTGAGAGAGTTTCTGACT--- 1061
 QY 733 GGCTTTGTACCAAGCACTTACCAACAGAGAAATTCACCTGTCATGAGTAAAGTTTTCAC 792
 Db 1062 GATTACTGTAAACAGTAAACCAACACCGGTGAATACAGTTGCTCAAAGTTCGATCTACTA 1121
 QY 793 CTGGAACGCGAGATGGCTACTATCTGATTCAGATGATCATCCAGCCTACTCATCTGTC 852
 Db 1122 TTCAGCGAGAAATCTCATATTACTTAATAACAAATTTATATACCATGCTGTATGTTGGTC 1181
 QY 853 ATCTGTCTCTGGTCTCTCTTCTGGATCAACATGGATGCTGCCCTCCCGTGTGGCGCTG 912
 Db 1182 ATTGTATCATGGTATCATCTTGGTGGATCAAGAGCAGTACCGGCGGAGTGTCACTG 1241
 QY 913 GCATCACACCGTGTCTACCAATGACCCAGAGCTGTGGCTCCCGGCGCTCTTTGGCT 972
 Db 1242 GGTGTCAACCCCTGCTGACCATGCGCCACCCAGAGCTCGGGCATAAACGCTCCCTGCCG 1301
 QY 973 AAGTGTCTTACGTGAAGCATGACATCTCGATGGCTGTGTGTCGCTCTTTGTGTTTC 1032
 Db 1302 CCCGTTTCTTATACGAAGCCATCGATGTGTGAGCAGCGGTGTGTCTGACGTTCTGTTTC 1361
 QY 1033 GCTGCTTCTGAGTATGCTGCCATAAATTTTGTTCG 1073
 Db 1362 GGGGCCCTGCTGAGTTCGCCCTGGTGAATGATCCG 1402

RESULT 14

US-09-949-016-16344
 ; Sequence 16344, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 16344
; LENGTH: 73519
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: missing
; LOCATION: (1)
; OTHER INFORMATION:
US-09949-016-16344

```

	Query Match	10.8%;	Score 176.8;	DB 4;	Length 73519;
	Best Local Similarity	87.7%;	Pred. No. 2.4e-41;		
	Matches 193;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	758	CAGGGAAATTCACCTGCATCGAGGTAAGTTTCACCTGGAAACGGCAGATGGGCTACTATC	817		
DB	42818	CAGGTAAATTCACCTGCATTCGATGAGCCCGTTCCACTCGAGCGGCAGATGGGTACTACC	42877		
QY	818	TGATTCAGATGTATACCCGACGCTACTCATCGTCAATCTGCTCGGGTCTCCTTCTCGA	877		
DB	42878	TGATTCAGATGTATATTCGCACGCTGCTCATTTGTCACTCTCATGGATCTCCTTCTCGA	42937		
QY	878	TCAACATGATGATGCTGCCCTGCCGCTGGCGCTGGGCATACCAACCGTGTCCACCATGA	937		
DB	42938	TCAACATGATGCTGCACCTGCTGTGTGGGCTAGGCATACCACTGTGCTCACCATGA	42997		
QY	938	CCACCCAGAGCTGTGCTCCCGGGCCTTTTGCCTAAGGT	977		
DB	42998	CCACCCAGAGCTCCGGCTCTCGAGCATCTCTGCCCAAGT	43037		

RESULT 15

```

US-09-949-016-11769
; Sequence 11769, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11769
; LENGTH: 105919
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(105919)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11769

```

	Query Match Best Local Similarity Matches 193; Conservative	10.8%; 87.7%; 0;	Score 176.8; DB 4; Pred. No. 3e-41; 0;	Length 105919; Indels 27; Gaps 0;
Qy	758	CAGGGAAATTCACCTGCATCGAGGTAAGATTTCACCTCGAAACGGCAGATGGGCTACTATC	817	
Db	75219	CAGGTAATTCACCTGCATTCAGGCCCGGTTCACCTCGAGCGGCAGATGGGTATTACTACC	75278	
Qy	818	TGATTTCAGATGTATACATCCCGCAGCCTACTCATGTCTATCCTCTGGGTCCTCCTTCTCGA	877	
Db	75279	TGATTTCAGATGTATATTCACGAGCCTGCTCATTTGTCATCTCTCATGTGGATCTCTTCTCGA	75338	
Qy	878	TCAACATGGATGTCGCCCTCTGCCCGGTGGGGCATACACACCGTGCTCACCATGA	937	

This Page Blank (uspto)

Qy	181	AGGATTCGGCCCAATTTTAAAGGCCCAACCCGTGAACGTGACCTGCAACATCTTCATCAAC	240
Db	181	AGGATTCGGCCCAATTTTAAAGGCCCAACCCGTGAACGTGACCTGCAACATCTTCATCAAC	240
Qy	241	AGTTTCAGCTCCGTCACCAAGACCACAAATGAGCTACCGGGTGAATGTCTTCTTGGCGCAA	300
Db	241	AGTTTCAGCTCCGTCACCAAGACCACAAATGAGCTACCGGGTGAATGTCTTCTTGGCGCAA	300
Qy	301	CAGTGGAAATGACCCACGCTGTCTTACCGAGAAATATCTGATGACTCTCTGGACCTCGAT	360
Db	301	CAGTGGAAATGACCCACGCTGTCTTACCGAGAAATATCTGATGACTCTCTGGACCTCGAT	360
Qy	361	CCCTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCC	420
Db	361	CCCTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCC	420
Qy	421	AAC TTCATGAGGTGACACGGCAACAAGTTACTGCGCATCTTCAAGAATGGGAATGTG	480
Db	421	AAC TTCATGAGGTGACACGGCAACAAGTTACTGCGCATCTTCAAGAATGGGAATGTG	480
Qy	481	CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCGCTGATGGACCTCAAGAACTTCCCC	540
Db	481	CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCGCTGATGGACCTCAAGAACTTCCCC	540
Qy	541	ATGGACATCCAGAGCTGCAAGATCGACTGTGAGAGCTCATCCATCTCTGACGCGCTCTG	600
Db	541	ATGGACATCCAGAGCTGCAAGATCGACTGTGAGAGCTCATCCATCTCTGACGCGCTCTG	600
Qy	601	CCATCTCTGTCACTTTTCAGTTGGCTCACCATGAAGACCTCGTGTTGTAGTGGCTGGAA	660
Db	578	-----TTGGGTCAACATGAAGACCTCGTGTTGTAGTGGCTGGAA	618
Qy	661	GATGCTCCTGTCTGTCAAAGTGGCTGAGGGGTGACTCTGCCCAAGTTTATCTTTCGGGGAT	720
Db	619	GATGCTCCTGTCTGTCAAAGTGGCTGAGGGGTGACTCTGCCCAAGTTTATCTTTCGGGGAT	678
Qy	721	GAGAAGATCTAGGCTGTGTACCAAGCACTACAACACAGGGAAATTCACCTGCATCGAG	780
Db	679	GAGAAGATCTAGGCTGTGTACCAAGCACTACAACACAGGGAAATTCACCTGCATCGAG	738
Qy	781	GTAAGGTTTCACTGGAACGGCAGTGGGTACTATCTGATTCAGATGTACATCCCGCAGC	840
Db	739	GTAAGGTTTCACTGGAACGGCAGTGGGTACTATCTGATTCAGATGTACATCCCGCAGC	798
Qy	841	CTACTCATCGTCATCCCTGCTCTGGGTCTCTTCTGGATCAACATGGATGCTGCCCTGCC	900
Db	799	CTACTCATCGTCATCCCTGCTCTGGGTCTCTTCTGGATCAACATGGATGCTGCCCTGCC	858
Qy	901	CGTGTGGGCTGGGCATCACACCGTGCTCAACATGACCCAGAGCTCTGGCTCCCGG	960
Db	859	CGTGTGGGCTGGGCATCACACCGTGCTCAACATGACCCAGAGCTCTGGCTCCCGG	918
Qy	961	GCCTCTTTGGCTTAAGGTGTCCTAGTGAAGGCAATCGACATCTGGATGGCTGTGTCTG	1020
Db	919	GCCTCTTTGGCTTAAGGTGTCCTAGTGAAGGCAATCGACATCTGGATGGCTGTGTCTG	978
Qy	1021	CTCTTTGTGTTTCGCTGCTGCTGGAGTATGCTGCCATAATTTTGTCTCGTCAGCAT	1080
Db	979	CTCTTTGTGTTTCGCTGCTGCTGGAGTATGCTGCCATAATTTTGTCTCGTCAGCAT	1038
Qy	1081	AAAGAAATTCATCGACTTCGAAGAGGCAGAGCGCCAAACGCTTGGAGGAAGATATCATC	1140
Db	1039	AAAGAAATTCATCGACTTCGAAGAGGCAGAGCGCCAAACGCTTGGAGGAAGATATCATC	1098
Qy	1141	CAAGAAAGTCGTTCTATATTCCTGGGCTATGCTGGGCCACGCTGCGAGGCAAGAGAT	1200
Db	1099	CAAGAAAGTCGTTCTATATTCCTGGGCTATGCTGGGCCACGCTGCGAGGCAAGAGAT	1158
Qy	1201	GGAGGTCCTAAGGAGTCTCGCATTTATAGTCCCAACCTCCAGCCCTCTCTTAAGG	1260
Db	1159	GGAGGTCCTAAGGAGTCTCGCATTTATAGTCCCAACCTCCAGCCCTCTCTTAAGG	1218

Qy	1261	GAAGGAGAAACACCGGGAAA	CTCTACGTGGAGCTGAGCCAGAGAAATTGACCACTCTCC	1322
Db	1219	GAAGGAGAAACACCGGGAAA	CTCTACGTGGAGCTGAGCCAGAGAAATTGACCACTCTCC	1278
Qy	1321	CGGGCTGTCTTCCCTTTT	CACCTTTCCCTCATCTTCAATATCTTCTACTGGGTTGCTATAAA	1380
Db	1279	CGGGCTGTCTTCCCTTTT	CACCTTTCCCTCATCTTCAATATCTTCTACTGGGTTGCTATAAA	1338
Qy	1381	GTGCTATGGTCAGAAGATAT	CACCAAGGCTCTGTGAATAGGGTGGAGCTTATAGAGTCCT	1440
Db	1339	GTGCTACGGTTCAGAAGATAT	CACCAAGGCTCTGTGAATAGGGTGGAGCTTATAGAGTCCT	1398
Qy	1441	GCTGCTGGCCTCTCTGCTT	CTCTCTGGTGGGCTTCTCCCTCAGTTAGACTCATTAGGG	1500
Db	1399	GCTGCTGGCCTCTCTGCTT	CTCTCTGGTGGGCTTCTCCCTCAGTTAGACTCATTAGGG	1458
Qy	1501	GTTTGGACAGTTCCTTCTT	CTTGATCTCCACCTCAGAACTTCAACTACCAGTCCCAAGCTAT	1560
Db	1459	GTTTGGACAGTTCCTTCTT	CTTGATCTCCACCTCAGAACTTCAACTACCAGTCCCAAGCTAT	1518
Qy	1561	GTGGGCTTATTTGCAATGGT	GCCAATGGTGGCTGTACTTATAAGATGGCTTATACCC	1620
Db	1519	GTGGGCTTATTTGCAATGGT	GCCAATGGTGGCTGTACTTATAAGATGGCTTATACCC	1578
Qy	1621	TA 1622		
Db	1579	TA 1580		
RESULT 3				
US-10-343-903-52				
; Sequence 52, Application US/10343903				
; Publication No. US20040224911A1				
; GENERAL INFORMATION:				
; APPLICANT: INCYTE GENOMICS INC.; YUE, Henry;				
; APPLICANT: THORNTON, Michael; RAMKUMAR, Jayalaxmi;				
; APPLICANT: TANG, Y. Tom; AZIMZAI, Yalda;				
; APPLICANT: BAUGHN, Mariah R.; YANG, Junming;				
; APPLICANT: YAO, Monique G.; LAL, Preeti G.;				
; APPLICANT: CHAWLIA, Narinder K.; GANDHI, Ameena R.;				
; APPLICANT: HAFALIA, April J.A.; NGUYEN, Dannel B.;				
; APPLICANT: ARVIZU, Chandra S.; ELLIOTT, Vicki S.;				
; APPLICANT: TRIBOULEY, Catherine M.; LU, Dyung Alina M.;				
; APPLICANT: XU, Yuming; REDDY, Roopa;				
; APPLICANT: HERNANDEZ, Roberto; BOROWSKY, Mark L.;				
; APPLICANT: LO, Terence P.; LU, Yan;				
; APPLICANT: POLICKY, Jennifer L.; GREENE, Barrie D.;				
; APPLICANT: SANJANWALA, Madhusudan M.; RAUMANN, Brigitte E.;				
; APPLICANT: BURFORD, Neil; ISON, Craig H.;				
; APPLICANT: LEE, Ernestine A.; DING, Li;				
; APPLICANT: DAS, Debopriya; KALLICK, Deborah A.;				
; APPLICANT: KHAN, Farah A.; SEILHAMER, Jeffrey J.;				
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS				
; FILE REFERENCES: FI-0183 USN				
; CURRENT APPLICATION NUMBER: US/10/343,903				
; CURRENT FILING DATE: 2003-02-03				
; PRIORITY APPLICATION NUMBER: PCT/US01/24217				
; PRIORITY FILING DATE: 2001-08-01				
; PRIORITY APPLICATION NUMBER: 60/231,434				
; PRIORITY FILING DATE: 2000-09-08				
; PRIORITY APPLICATION NUMBER: 60/230,067				
; PRIORITY FILING DATE: 2000-08-31				
; PRIORITY APPLICATION NUMBER: 60/228,140				
; PRIORITY FILING DATE: 2000-08-25				
; PRIORITY APPLICATION NUMBER: 60/226,410				
; PRIORITY FILING DATE: 2000-08-18				
; PRIORITY APPLICATION NUMBER: 60/224,456				
; PRIORITY FILING DATE: 2000-08-10				
; PRIORITY APPLICATION NUMBER: 60/223,269				
; PRIORITY FILING DATE: 2000-08-03				
; NUMBER OF SEQ ID NOS: 60				

Query Match	49.3%	Score 807.8	DB 14	Length 993
Best Local Similarity	90.0%	Pred. No. 6.7e-249		
Matches 931	Conservative 0	Mismatches 2	Indels 102	Gaps 2
QY	1	ATGACAACTCTTGTTCTCTGCAACCTCTCCTTCCCTTCTCTCTGACCCCTGCGAGGCGAG	60	
DB	1	ATGACAACTCTTGTTCTCTGCAACCTCTCCTTCCCTTCTCTCTGACCCCTGCGAGGCGAG	60	
QY	61	GTCTCTCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAACTCTGGAACCAAGGGGTCCCGAG	120	
DB	61	GTCTCTCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAACTCTGGAACCAAGGGGTCCCGAG	120	
QY	121	CCCATGTCCCTCTGATTTCTTAGACAAACTTATGGGGCGAAACATCTGGATATGATGCC	180	
DB	121	CCCATGTCCCTCTGATTTCTTAGACAAACTTATGGGGCGAAACATCTGGATATGATGCC	180	
QY	181	AGGATTCGGGCCCAATTTTAAAGGCCCAACCGTGAACCTGACCTGCAACATCTTCATCAAC	240	
DB	181	AGGATTCGGGCCCAATTTTAAAGGCCCAACCGTGAACCTGACCTGCAACATCTTCATCAAC	240	
QY	241	AGTTTTCAGCTCCGTCACCAAGACCAAAAT	269	
DB	241	AGTTTTCAGCTCCATCACCAAGACCAAAATGGCTTGTCTGGGCCCTGGGAATGGCAATGTT	300	
QY	270	-----GGACTACCGGGTGAATGTCTTCTTCGGCGAA	300	
DB	301	TCTGAAGGGGCCATATCTGCACCCCTCCAGGACTACCGGGTGAATGTCTTCTTCGGCGAA	360	
QY	301	CAGTGGAAATGACCAACGCTGTCTACCGAGAATATCTGATGACTCTCTGGACCTCGAT	360	
DB	361	CAGTGGAAATGACCAACGCTGTCTACCGAGAATATCTGATGACTCTCTGGACCTCGAT	420	
QY	361	CCCTCCATGTGGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCC	420	
DB	421	CCCTCCATGTGGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCC	480	
QY	421	AACCTCCATGAGGTGACCAAGTAACTCTGCGCATCTTTCAAGAAATGGGAATGTG	480	
DB	481	AACCTCCATGAGGTGACCAAGTAACTCTGCGCATCTTTCAAGAAATGGGAATGTG	540	
QY	481	CTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC	540	
DB	541	CTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC	600	
QY	541	ATGGACATCCAGACGTGCACGATGCAGCTTGAGAGCTCATCCATCTCTGCAGCCCTCTG	600	
DB	601	ATGGACATCCAGACGTGCACGATGCAGCTTGAGAGCT-----	637	
QY	601	CCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAAGACCTCTGTTTGTAGTGGCTGAA	660	
DB	638	-----TTGGCTACACCATGAAAGACCTCTGTTTGTAGTGGCTGAA	678	
QY	661	GATGCTCTCTGTCCAAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTCGGGGAT	720	
DB	679	GATGCTCTCTGTCCAAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTCGGGGAT	738	
QY	721	GAGAAGGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGAATTCACCTGCTATCGAG	780	
DB	739	GAGAAGGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGAATTCACCTGCTATCGAG	798	
QY	781	GTAAAGTTTCACTGGAAACGCGAGATGGCTACTATCTGATTCAGATGTATACATCCCCAGC	840	
DB	799	GTAAAGTTTCACTGGAAACGCGAGATGGCTACTATCTGATTCAGATGTATACATCCCCAGC	858	
QY	841	CTACTCATGTCATCTCTGCTCTGGGTCTCTTCTTGGATCAACATGGATGCTGCCCTGCC	900	
DB	859	CTACTCATGTCATCTCTGCTCTGGGTCTCTTCTTGGATCAACATGGATGCTGCCCTGCC	918	
QY	901	CGTGTGGGCTTGGGCATCACCAACGCTGTACACCATGACCAACCCAGAGCTCTGGCTCCCGG	960	
DB	919	CGTGTGGGCTTGGGCATCACCAACGCTGTACACCATGACCAACCCAGAGCTCTGGCTCCCGG	978	
QY	961	GCCTCTTTTGCTTAAG	975	

```

Db          979 GCCTCTTGGCCTAAAG 993
|||||
RESULT 5
US-10-757-262-9
; Sequence 9, Application US/10757262
; Publication NO. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 14351, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84360, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MPI03-007PIRNOWNM1
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (297)...(1646)
US-10-757-262-9

Query Match      40.0%; Score 656; DB 18; Length 1715;
Best Local Similarity 71.9%; Pred. No. 7.7e-200;
Matches 955; Conservative 0; Mismatches 315; Indels 57; Gaps 5;

Qy      89 AGGAAGTCAAACTGGAAACCAAGGGGTCCAGGCCCATGTCCCTCTGATTTCTTAGACA 148
Db      367 AGGAGGCTGAAGCTGCTCGCTCCGCAACCAAGCCTATGTCAACCTCGGATTTCTGGATA 426

Qy      149 AACTTATGGGGCGAACAATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC 208
Db      427 AGCTAATGGGGAGAACCTCCGGATATGATGCCAGGATCAGGCCCAATTTTAAAGGTCCCC 486

Qy      209 CCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTTCAGTCCGTCACCAAGCACCAACAA 268
Db      487 CAGTGAACGTGACCTGCAACATTTTCATCAACAGCTTTGGTTCATCTGTGAGACAACCA 546

Qy      269 TGGACTACCGGGGTGAATGTCTTCTTCGGGCAACAGATGGAATGATCCACGCGCTGCTTACC 328

```

Db 547 TGGACTATAGGTCAACATCTTCTCGGAGCAATGGAACGACCCCGCTGGCCTATA 606
 QY 329 GAGAAATATCTGATGACTCTCTGGAGCTCGATCCCTCCATGCTGGACTCTATCTGAAGC 388
 Db 607 ATGAATACCTGACGACTCTCTGGAGCTGAGCCATCCATCTGGACTCACTCGGAAC 666
 QY 389 CAGACCTCTCTTGTGTAATGAGAAAGGGCCCACTTCCATGAGGTGACCAAGCAACA 448
 Db 667 CTGACCTGTTCTTGGCAAGAGAGGGGCCACTTCCATGAGATCACCACAGACAACA 726
 QY 449 AGTTACTGGGCATCTTCAAGATGGAATGTGTGTACAGCATCAGGCTGACCTCATTT 508
 Db 727 AATTGTAAGGATCTCCCGAATGGAATGTCTCTACAGCATCAGATCAACCTGACAC 786
 QY 509 TGTCTGCTGATGAGACCTCAAGAACTTCCCATGACATCCAGAGCTGCAGATCAGC 568
 Db 787 TGGCCTGCCCCATGGACTTGAAGATTTCCCATGATGTCAGACATGATATCAAC 846
 QY 569 TTGAGAGCTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTCAGTTGGCTACA 628
 Db 847 TGAAGAGCT-----TTGGATATA 864
 QY 629 CAGTAAAGACCTGTTTGTAGTGGCTGGAAGATGCTCTGCTGCCAAGTGGCTGAGG 888
 Db 865 CGATGAATGACCTCATCTTGTAGTGGCAGGAACAGG---GAGCGGTGCAGGTAGCAGATG 921
 QY 689 GGTGACTCTGCCCAAGTTTATCTTGGGATGAGAGGATCTAGGCTGTGTACCAAGC 748
 Db 922 GACTAATCTGCCCAAGTTTATCTTGAAGAGAGGAGACTTGAATATCTGACCAAGC 981
 QY 749 ACTACACACAGGAAATTCACCTGATCGAGTAAAGTTTCACTTGGAAAGCAGATGG 808
 Db 982 ACTACACACAGTAAATTCACCTGATTCAGGCGCGTTCACCTGGAGCGCAGATGG 1041
 QY 809 GCTACTATCTGATTCAGATGATCATCCAGGCTTCTCATGTCATCTCTGCTGCTGCT 868
 Db 1042 GTTACTACCTGATTCAGATGATATTCAGGCTGCTCATTTGTCATCTCTCATGATCT 1101
 QY 869 CTTCTGGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
 Db 1102 CTTCTGGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
 QY 929 TCACATGACCAACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
 Db 1162 TCACATGACCAACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221
 QY 989 AGCAATCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
 Db 1222 AAGCAATTCGATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1281
 QY 1049 ATGCTGCCATTAATTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108
 Db 1282 ATGCTGCCATTAATTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1341
 QY 1109 AGAGGCGCAACGCTTGGAGGAGATATCATCAAGAAAGTCTGTTTCTATTTCCGCTGCT 1168
 Db 1342 GGAGACATCACAGGAGGATGAAGCT-----GGAGAGCGCTTTAACTTCTGCTGCT 1395
 QY 1169 ATGGCTTGGGCC---ACTGCTGAGGAGAGATGAGGCTCAATGGAAGGTTCT---TG 1222
 Db 1396 ATGGGATGGGCCGAGCTGCTACAGGCAAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCT 1455
 QY 1223 GCATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGAGAGGAGAAACACCGCGAAGC 1282
 Db 1456 ACAGTAACACCAACACCCCTCTCTGACCATCTAAGTCCCGAGAGAGATGCGAAGC 1515
 QY 1283 TCTACGTGGACTGAGCAAGAGAAATGACACCATCTCCCGGGCTGCTTCCCTTTTCACTT 1342
 Db 1516 TCTTCAATCAGAGGGCCAGAGATCGACAAATATATCCGCAATGCTTCCCGCTGCTGCT 1575
 QY 1343 TCCTCATCTTCAATATCTTCTACTGGGTGCTATTAAGTGTCTATGCTGAGAGATATCC 1402

Db 1576 TCCTCATTTTCAACATGTTTCTACTGATCATCTACAAGATTGTCCTAGAGAGACGTCC 1635
 QY 1403 ACCA 1406
 Db 1636 ACA 1639
 RESULT 6
 US-10-205-194-20
 ; Sequence 20, Application US/10205194
 ; Publication No. US20030134301A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warner-Lambert Company
 ; APPLICANT: Lee, Kevin
 ; APPLICANT: Dixon, Alistair
 ; APPLICANT: Brooksbank, Robert
 ; APPLICANT: Pincock, Robert
 ; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
 ; FILE REFERENCE: WL-A-018201
 ; CURRENT APPLICATION NUMBER: US/10/205,194
 ; CURRENT FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: GB 0118354.0
 ; PRIOR FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 177
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 20
 ; LENGTH: 3865
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Neonatal glycine receptor
 US-10-205-194-20
 Query Match 39.9%; Score 654.8; DB 15; Length 3865;
 Best Local Similarity 70.8%; Pred. No. 2.9e-199;
 Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
 QY 84 AAAAGAGGAGTCAAAATCTGGAACCAAGGGTCCAGAGCCATGTCCTCTGATTTCTT 143
 Db 1256 AGACCATGACTCCAGTCTGGAACACATCCCTCGCAGACCCCTGCTCTCAGATTTCTT 1315
 QY 144 AGACAACTTATGGGGGGAACATCTGGATATATGATGCGAGGATTCGCGCCCAATTTTAAAG 203
 Db 1316 GGATAAACTTAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAG 1375
 QY 204 CCCACCGTGAAGTCACTGACCTGCAACATCTTCAACAGTTTTCAGCTTTCAGTTCGAC 263
 Db 1376 TCCTCAGTAAACGTTTACTTGCATATTTTATCAAGTTTGGATTCGTCACAGAAC 1435
 QY 264 CACAATGGAATACCGGGTGAATGTCTTCTTGGGCAACAGTGAATGACCCAGCCTGTC 323
 Db 1436 CACCATGGAATACCGAGTGAACATTTTCTGAGACAGCAGTGAACGATTCAGCGCTGGC 1495
 QY 324 CTACCGAGAAATATCTGATGACTCTCTGAGACCTGATTCCTCTCATGCTGAGACTCTATCTG 383
 Db 1496 ATACGTGAGTACCCAGATGATTCCTGATTTGGATTCATCGATGCTGATTCGATTTG 1555
 QY 384 GAAGCAGACCTCTCTTCTTGTAAATGAGAAAGGGGCAACTTCCATGAGTGAACACGGA 443
 Db 1556 GAACACAGATTTGTTCTTGTGCAATGAGAAAGAGCAATTTCCATGATGTCACCACTGA 1615
 QY 444 CAACAAGTTACTGCGCATCTTCAAGAAATGGGAATGTGTGTATCAGCATCAGGCTGACCT 503
 Db 1616 TAACAAGTTGCTGCGGATTTCCAAAATGCGAAAGTGTCTACAGTATTAGACTCACTT 1675
 QY 504 CATTTTGTCTGCTGATGAGCTCAAGAACTTCCCATGGAATCCAGACGTCAGAT 563
 Db 1676 GACTTTATCTGCTCCAGCTGAGAACTTTTCAATGGATGTCCAGACCTGTCAAT 1735
 QY 564 GCAGCTTGAAGAGCTATCCATCTCTGAGCCCTCTGCACTCTCTGCTCAGTTTCACTT 623
 Db 1736 GCAGCTGAGAGGTT-----TTGG 1753

QY 624 CTACACCATGAAGACCTCTGTTGAGTGGCTGGAAGATGCTCCTGCTGCTCAAGTGGC 683
 Db 1754 GTACACCATGAAGACCTGATATTTGAGTGGTAAAGTATGGTGC---CAGTACAAGTTGC 1810
 QY 684 TGAGGGGCTGACTCTGCCAGTTTATCTTGGGGATGAGAAGATCTAGGCTGTTGTAC 743
 Db 1811 TGAAGGACTCACCTGCTCAGTTTATTTGAAAGAGAGAAAGAACTTGGCTATTGCAC 1870
 QY 744 CAAGCACTACACACAGGGGAAATTCACCTGCATCGAGGTAAAGTTTCACTCGAAGCGCA 803
 Db 1871 AAAGCACTTACACACTGGCAAGTTTACCTGCATTTGAGGTCAAGTTTCACTCGAAGCGCA 1930
 QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCCGAGCTTACTCATCTGTCATCTGTCCTG 863
 Db 1931 GATGGGCTATTATTGATCCAGATGTATATCCCGAGCTTCTGATAGTCAATTTTGTCCCTG 1990
 QY 864 GGTCTCCTCTCGGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
 Db 1991 GGTCTCCTCTCGGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2050
 QY 924 CGTGCTCCTCTCGGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
 Db 2051 CGTGCTCCTCTCGGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2110
 QY 984 CGTGAAGGCAATCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1043
 Db 2111 TGTGAAGCAATTCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2170
 QY 1044 GGAGTATGCTGCCATAAATTTGTTTCTCGTACGATTAAGAAATTCATAGACTTCGAAAG 1103
 Db 2171 GGATATGACAGAGTGAATTTGTTTCTCGAGCAATAAGAAATTCATAGACTTCGAAAG 2230
 QY 1104 AAGGAGAGCGCAACGCTTGGAGGAAGATATCATCAAGAAAGTGGTTTCTATTTCGG 1163
 Db 2231 ACACAGAGAGCGCAATTAAGCAAGAAAGTGTACTCGTGAAGCGTTTAACTTCAG 2290
 QY 1164 TGCTATGCTTGGGCGACCTGCTGAGGCAAGAGATGAGGTCCATGGAAGGTTCTGG 1223
 Db 2291 CGGTATGAGGATGGGTGACCTGCTCCTCAAGTGAAGATGCTACAGCTGTCAAGGCTAC--- 2347
 QY 1224 CATTTATAGTCCCACTCCAGCCCTCTTCAAGGAGGAGAAACACACGCGGAACT 1283
 Db 2348 -----ACCTGCCAACCCACTTCGGCAACCCCAAGATGCTATCAAGAGAA 2401
 QY 1284 CTAGTGGATGACCAAGAGAAATGACACCATCTCCGGGCTGTCTCCCTTTCACTTT 1343
 Db 2402 GTTTGAGTACGGGCAAAAGAAATCGACACCATCTCGAGTGCCTTCCCACTGGCCTT 2461
 QY 1344 CCTCATCTTCAATATCTTCTACTGGGTTGTCTATAAGTGTCTATGTCAGAAATATCCA 1403
 Db 2462 CCTCATTTTCAACATCTTTTACTGGATCACATCAAGATCATTCGGCATGAAGATGTCCA 2521
 QY 1404 CCAG 1407
 Db 2522 CAAG 2525

RESULT 7

US-10-486-706-235
 ; Sequence 235, Application US/10486706
 ; Publication No. US20050071088A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LANDFIELD, PHILIP W.
 ; APPLICANT: BLALOCK, ERIC M.
 ; APPLICANT: CHEN, KUEY-CHU
 ; APPLICANT: FOSTER, THOMAS C.
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR
 ; TITLE OF INVENTION: BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT
 ; FILE REFERENCE: 50229-426
 ; CURRENT APPLICATION NUMBER: US/10/486, 706
 ; CURRENT FILING DATE: 2004-02-13
 ; PRIOR APPLICATION NUMBER: PCT/US02/25607
 ; PRIOR FILING DATE: 2002-08-13

; PRIOR APPLICATION NUMBER: US 60/311,343
 ; PRIOR FILING DATE: 2001-08-13
 ; NUMBER OF SEQ ID NOS: 461
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 235
 ; LENGTH: 3865
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; US-10-486-706-235
 Query Match 39.9%; Score 654.8; DB 19; Length 3865;
 Best Local Similarity 70.8%; Pred. No. 2.9e-199;
 Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
 QY 84 AAAAGAGGAGTCAAAATCTGGAAACCAAGGGGTCCAGCCCATGTCCCTCTGATTTCTT 143
 Db 1256 AGACCATGACTCCAGGTCTGGAAACATCCCTCGCAGACCTCTCTCTTCAGATTTCTT 1315
 QY 144 AGCAAACTTATGGGGCGAACAATCTGGATATGATGCGAGATTCGGGCCCAATTTTAAAG 203
 Db 1316 GGATAAACTAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAG 1375
 QY 204 CCACCCGCTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTACCAAGAC 263
 Db 1376 TCCTCAGTAAAGCTTACTTGCATATTTTATCAACAGTTTGGATCGGTCAAGAAC 1435
 QY 264 CAAATGGACTACCGGGTGAATGTCTTCTGGGCAACAGTGGAAATGACCCACGCTGTC 323
 Db 1436 CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGGAAACGATTCACGGCTGGC 1495
 QY 324 CTACCGAAGATATCTGATGACTCTCTGGACCTCGATCCCTCATCTGCTGACTCTATCTG 383
 Db 1496 ATACAGTGAATACCCAGATGATTCCTGATTTGGATTCATCGATGCTGGATTCGATTTG 1555
 QY 384 GAAGCAGACCTCTCTTCTGCTAAAGAGAAAGGGGCAACTTCCATGAGTGAACCGGA 443
 Db 1556 GAACCAAGATTTGTTCTTGGCAATGAGAAAGAGCAATTTCCATGATGTACCACTGA 1615
 QY 444 CAACAGTTACTGCGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATCAGGCTGACCT 503
 Db 1616 TAACAGTTGCTGGGATTTCCAAATTCGAAAGTGTCTACAGATTTAGACTCACTT 1675
 QY 504 CATTTGCTGCTGATGAGACCTCAAGAACTTCCCATGGACATCCAGCTGACAGAT 563
 Db 1676 GACTTTATCTGTCATGAGACCTGAAGAACTTTTCAATGGATGTCCAGACCTGTCAAT 1735
 QY 564 GCAGCTTGAAGAGCTCATCTACTCTGCGAGCCCTCTGCCATCTCTCTGCTCACTTTT 623
 Db 1736 GCAGCTGGAGAGTT-----TTGG 1753
 QY 624 CTACACCATGAAGACCTCTGTTGAGTGGCTGGAAGATGCTCTCTGCTGTCCAAAGTGC 683
 Db 1754 GTACACCATGAATGACTGATATTTGAGTGGTAAAGTATGATGCTC---CAGTACAAGTGC 1810
 QY 684 TGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGATCTAGGCTGTTGTAC 743
 Db 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAGAGAGAACTTGGCTATTGTCAC 1870
 QY 744 CAAGCACTACACACAGGGGAAATTCACCTGCATCGAGGTAAAGTTTCACTCGAAGCGCA 803
 Db 1871 AAAGCACTTACACACTGGCAAGTTTACCTGCTGATTTAGGTCAAGTTTCACTCGAAGCGCA 1930
 QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCCGAGCTTACTCATCTGCTGCTGCTGCTG 863
 Db 1931 GATGGCTATTATTGATCCAGATGTATATCCCGAGCTGCTGATAGTCAATTTTGTCTG 1990
 QY 864 GGTCTCCTCTCGGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
 Db 1991 GGTCTCCTTTTGGATAAAACATGGATGACGCTCTCTGCGAGGTTGCTTGGCATCAAC 2050
 QY 924 CGTGCTCACCATGACCCAGAGCTCTGCTGCTCCCGGCTCTTTGCTTAAAGTGTCTTA 983
 Db 2051 CGTCTGCAATGACTTACCAGAGTTTCAGGTTCCAGGGCATCTCTGCCAAAGGCTCTCAT 2110

920 CCACCGTGTACCATGACACCCAGAGCTCTGGCTCCGGGCTCTTTGGCTTAAGGTG 979
 1292 CCACCGTGTAAACGATGACTACACAGATTTCAGGATTCAGGCTTCTTGGCCAAAGTTT 1351
 980 CTTAGCTGAAGCAATCGACATCTGATGGCTGTGTCTGTCTTGTGTGGCTGCCT 1039
 1352 CATATGTCAAAGCTATTGATTTTGGATGGCAGTATGCCTCTTTTGTGTTCAGCAC 1411
 1040 TGGTGGAGTATGTGCGCAATATTTTGTCTCTGTCAGATTAAGAAATTCATACGACTTC 1099
 1412 TTTGGAGTATGAGCTGTAAATTTTGTATCAAGACAACAAGAACTTCTGAGATTTC 1471
 1100 GAAGAAGGAGAGCGCCACGCTTGGAGGAAGATATCATCCAAAGAAAGTCGTTCTATT 1159
 1472 GACGAAGAGAGAG-----AATAAGGATGATGAGGTAAAGGAAAGCGGATTCAGCT 1522
 1160 TCCGTGGCTATGGCTTGGGCCATCTGCTGAGCAAGAGATGAGGTCCAAATGGAAGTT 1219
 1523 TCACAGCTATGGAATGGGACCATGTCTACAAGCAAAAGGATGCATCACTCCAAAGGGCC 1582
 1220 CTGGCATTTATAGTCCCCCACTCCAGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGA 1279
 1583 C-----CAACCCCTCTCCAGTAATGCCAAAGTCTGATGAATGAGGA 1630
 1280 AACTCTAGCTGAGTACGAGCAAGAGAAATGACACCATCTCCGGGCTGTCTCCCTTTCA 1339
 1631 AGGTCTTTATCGACCGGGCAAGAGATTGATACCATCTCCCGAGCTGTCTCCATTAG 1690
 1340 CTTTCTCATCTTCAATATCTTCTACTGGTGTCTTATAAAGTGTATGGTCAGAAAGATA 1399
 1691 CTTTTTGTATTTTAAATATTTCTACTGGGTATCTATAAAATTTCTAGGCATGAGGATA 1750
 1400 TCCACAG 1407
 1751 TCCATCAG 1758

RESULT 10
 US-10-075-846-5
 ; Sequence 5, Application US/10075846
 ; Publication No. US20030032608A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT B
 ; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF
 ; FILE REFERENCE: D0079 NP
 ; CURRENT APPLICATION NUMBER: US/10/075,846
 ; CURRENT FILING DATE: 2002-02-13
 ; PRIOR APPLICATION NUMBER: US 60/269,535
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 679
 ; TYPE: DNA
 ; ORGANISM: homo sapiens

US-10-075-846-5
 Query Match 29.3%; Score 480.6; DB 14; Length 679;
 Best Local Similarity 99.2%; Pred. No. 1.6e-143;
 Matches 483; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

1118 AACGCTTGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCTGGCTATGGCTTGG 1177
 193 AATTCAGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCTGGCTATGGCTTGG 252
 1178 GCCACTGCTTCGAGCAAGAGATGGAGGTCCAATGGAAGTTCTGGCAATTTATAGTCCCC 1237
 253 GCCACTGCTTCGAGCAAGAGATGGAGGTCCAATGGAAGTTCTGGCAATTTATAGTCCCC 312
 1238 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCGCGGAACTCTACGTGACTGAG 1297
 313 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCGCGGAACTCTACGTGACTGAG 372

1298 CCAAGAGAATTGACACCATCTCCCGGCTGTCTTCCCTTTTCACTTTCTCTCAATA 1357
 373 CCAAGAGAATTGACACCATCTCCCGGCTGTCTTCCCTTTTCACTTTCTCTCAATA 432
 1358 TTTTCTACTGGGTGTCTATAAAGTGTATGTCAGAAAGATATCCACAGGCTCTGTGAA 1417
 433 TTTTCTACTGGGTGTCTATAAAGTGTATGTCAGAAAGATATCCACAGGCTCTGTGAA 492
 1418 TAGGTGGAGACTATAGAGTCTGCTGCTGGCTCTCTGCTTCTCTGCTGGGCTTCT 1477
 493 TAGGTGGAGACTATAGAGTCTGCTGCTGGCTCTCTGCTTCTCTGCTGGGCTTCT 552
 1478 CCTCAGTTTAGACTCCATTAGGGGTTTGACAGTTCCTTCTCTGATCTCCCACTCAGAACT 1537
 553 CCTCAGTTTAGACTCCATTAGGGGTTTGACAGTTCCTTCTCTGATCTCCCACTCAGAACT 612
 1538 TCAACTACAGTCCCAAGCTATGTGGGCTATATGCAATGTCGCAATGGTGGCTGTAC 1597
 613 TCAACTACAGTCCCAAGCTATGTGGGCTATATGCAATGTCGCAATGGTGGCTGTAC 672
 1598 TTATAAA 1604
 673 TTATAAA 679

RESULT 11
 US-09-802-668-49/c
 ; Sequence 49, Application US/09802668
 ; Publication No. US20030190714A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wood, Linda
 ; APPLICANT: Vogel, Gabriel
 ; APPLICANT: Karnovsky, Alla
 ; APPLICANT: Linske-O'Connell, Lisa I.
 ; APPLICANT: Wang, Jun
 ; APPLICANT: Liu, Derong
 ; TITLE OF INVENTION: Human Ion Channels
 ; FILE REFERENCE: 00069US1
 ; CURRENT APPLICATION NUMBER: US/09/802,668
 ; CURRENT FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/188,517
 ; PRIOR FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 117
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 49
 ; LENGTH: 554
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-09-802-668-49
 Query Match 18.2%; Score 299; DB 10; Length 554;
 Best Local Similarity 98.1%; Pred. No. 4.9e-85;
 Matches 313; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

1118 AACGCTTGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCTGGCTATGGCTTGG 1177
 319 AATTCAGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCTGGCTATGGCTTGG 260
 1178 GCCACTGCTTCGAGCAAGAGATGGAGGTCCAATGGAAGTTCTGGCAATTTATAGTCCCC 1237
 259 GCCACTGCTTCGAGCAAGAGATGGAGGTCCAATGGAAGTTCTGGCAATTTATAGTCCCC 200
 1238 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCGCGGAACTCTACGTGACTGAG 1297
 199 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCGCGGAACTCTACGTGACTGAG 140
 1298 CCAAGAGAATTGACACCATCTCCCGGCTGTCTTCCCTTTTCACTTTCTCTCAATA 1357
 139 CCAAGAGAATTGACACCATCTCCCGGCTGTCTTCCCTTTTCACTTTCTCTCAATA 80
 1358 TTTTCTACTGGGTGTCTATAAAGTGTATGTCAGAAAGATATCCACAGGCTCTGTGTA 1416


```

Db 79 TCTTCTACTGGTGTCTATAAAGTCTACGGTCTAGAGATATCCACAGGCTTCTGTGA 20
QY 1417 ATAGGGTGGAGCTATAGA 1435
Db 19 ATAGGGTGGAGCTATAGA 1
RESULT 12
US-10-243-475-49/c
; Sequence 49, Application US/10243475
; Publication No. US20030194720A1
; GENERAL INFORMATION:
; APPLICANT: Roberds, Steven L.
; APPLICANT: Benjamin, Christopher W.
; APPLICANT: Karnovsky, Alla M.
; APPLICANT: Ruble, Cara L.
; TITLE OF INVENTION: Human Ion Channels
; FILE REFERENCE: PHRM0039-100
; CURRENT APPLICATION NUMBER: US/10/243,475
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/403,254
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 60/318,733
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-243-475-49
Query Match 18.2%; Score 299; DB 16; Length 554;
Best Local Similarity 98.1%; Pred. No. 4.9e-85;
Matches 313; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1118 AACGCTTGGAGGATATCATCAAGAAAGTCTTCTATTTCCGTGGCTATGGCTTGG 1177
Db 319 AATTCAGGAGGAGATATCATCAAGAAAGTCTTCTATTTCCGTGGCTATGGCTTGG 260
QY 1178 GCCACTGCCCTCAGGCAAGAGATGGAGTCCATGGAAGTCTCTGGCAATTTATAGTCCCC 1237
Db 259 GCCACTGCCCTCAGGCAAGAGATGGAGTCCATGGAAGTCTCTGGCAATTTATAGTCCCC 200
QY 1238 AACCTCCAGCCCTCTTCTTAAGGGAAGGAGAAACACGCGGAACTCTACGTGGACTGAG 1297
Db 199 AACCTCCAGCCCTCTTCTTAAGGGAAGGAGAAACACGCGGAACTCTACGTGGACTGAG 140
QY 1298 CCAAGAGATTTGACCATCTCCCGGCTGTCTTCCCTTTCACCTTTCCTCATCTTCAATA 1357
Db 139 CCAAGAGATTTGACCATCTCCCGGCTGTCTTCCCTTTCACCTTTCCTCATCTTCAATA 80
QY 1358 TCTTCTACTGGTGTCTATAAAGTCTATGGTCTAGAGATATCCACAGGCTCTGTGTA 1416
Db 79 TCTTCTACTGGTGTCTATAAAGTCTATGGTCTAGAGATATCCACAGGCTTCTGTGA 20
QY 1417 ATAGGGTGGAGCTATAGA 1435
Db 19 ATAGGGTGGAGCTATAGA 1
RESULT 13
US-10-305-720-271
; Sequence 271, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30

```

```

; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 271
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 1695770
US-10-305-720-271
Query Match 15.4%; Score 253; DB 17; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.1e-70;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1004 GGATGGCTGTGTCTGTCTCTTTGTGTTCGCTGCTTGTGGAGTATGTCGCAATAAAT 1063
Db 1 GGATGGCTGTGTCTGTCTCTTTGTGTTCGCTGCTTGTGGAGTATGTCGCAATAAAT 60
QY 1064 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAGGCGAGCGCAACGCT 1123
Db 61 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAGGCGAGCGCAACGCT 120
QY 1124 TGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCGTGGCTATGCTGGGCACT 1183
Db 121 TGGAGGAAGATATCATCCAAAGAAAGTCGTTTCTATTTCCGTGGCTATGCTGGGCACT 180
QY 1184 GCCTGAGGCAAGAGATGGAGTCCAAATGGAAGTTCTGGCAATTTATAGTCCCAACCTC 1243
Db 181 GCCTGAGGCAAGAGATGGAGTCCAAATGGAAGTTCTGGCAATTTATAGTCCCAACCTC 240
QY 1244 CAGCCCTCTTCT 1256
Db 241 CAGCCCTCTTCT 253
RESULT 14
US-10-172-118-529
; Sequence 529, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 529
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_000824
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-529
Query Match 14.2%; Score 232.6; DB 17; Length 2106;
Best Local Similarity 56.6%; Pred. No. 2.6e-63;
Matches 504; Conservative 0; Mismatches 339; Indels 48; Gaps 2;
QY 170 GATATGATCCAGGATTCGGCCCAATTTTAAAGGCCACCGCTGAACGTGACCTGCAACA 229
Db 273 GTTATGATCCAGGATAAGCAACTTCNAAGGCATTCCTGTTGATAGTAGTCAACA 332

```

```

QY 230 TCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCAATGGACTACCGGTGAATGTCT 289
Db 333 TTTTATTAACAGTTTGGATCCATTCAAGAAACAATGGACTATAGATTAAATCT 392
QY 290 TCTTGGGCAACAGTGAATGACCCAGCGCTGTCCTACCGAGATAT-----CCTGATG 343
Db 393 TCTGAGACAAAATGGAATGACCCAGCGCTGAAGCTCCCAAGTATTTAGGGGTTCAG 452
QY 344 ACTCTCTGGACCTCGATCCCTCATCTGGACTCTATCTGGAAGCAGACCTCTCTTTTG 403
Db 453 ATGCACTGACAGTGGATCCCAATTTTCATGATGTGACCCAGGAAACATCTCTTTATTT 512
QY 404 CTAAATGAGAAAGGGGCAACTTCCATGAGTGACCGACCAACAAAGTACTGCGCATCT 463
Db 513 CAATGAAAGAAAGTGCCTTTTCATGATGTGACCCAGGAAACATCTCTCTTTATTT 572
QY 464 TCAAGATGGGAATGTGCTGACAGATCAGGCTGACCCCTCATTTTGTCTGCTGATGG 523
Db 573 TCGTGTGAGATGTCTTGTGAGATGAGGTATCTATTCTTTTCATGCTTTCATGCTTGG 632
QY 524 ACCTCAAGAACTTCCCATGGACATCCAGAGCTGACAGTGCAGCTTGAGAGCTCATCCA 583
Db 633 ACTTGACATGTTTCCATGGATACAAAGCTTGCAAGTGTCAACTGGAGAGCT----- 686
QY 584 TACTCTGCAGCCCTCTGCCATCTCTCTCTTTCAGTTGGCTACACCATGAAGACCTCG 643
Db 687 -----TTGTTTACAACTGATTTAC 710
QY 644 TGTTTGGTGGTGAAGATGCTCTCTGCTGTCCAAAGTGGCTGAGGGGCTGACTCTGCCCC 703
Db 711 GATTTATCTGGCAGTCAAGAGATCTCTGCTCAATAGAAAATTTGCTTGCCTCAATTTG 770
QY 704 AGTTTATCTTGGGGATGAGAAGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGA 763
Db 771 ATATCAAAAGGAAGATATTGAATATGTTAACTGTACAAAATATATAAGGCACGGCT 830
QY 764 AATTACCTGCTGAGGTAAAGTTTCACTTGAACGGCAGATGGGCTACTATCTGATTC 823
Db 831 ACTACACATGCTGGAAAGTCTCTTCCACCTGAGGAGGAGGTCGGCTTTTACATGATG 890
QY 824 AGATGCTGCTGCCCTCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 883
Db 891 GGGTCTACGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
QY 884 TGGATGCTGCCCTCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 943
Db 951 CGGACGCGAGTGTGCGCAGAGTGCCTTGGTATCTTCTCAGTCTCAGTCTGCTGCTGCT 1010
QY 944 AGAGCTCTGGCTCCCGGGCTCTTTGGCTTAAAGTGTCTTGAAGGCAATCGACATCT 1003
Db 1011 AGTGCAAACTTGGCTGAGCTTCCCAAGTTTCTCTATGTGAAGGCTCTTGTATGTTT 1070
QY 1004 GATGCTGTGCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054
Db 1071 GGCTTATGCTGCTCTCTTGTGGTGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121

```

RESULT 15

```

US-10-342-887-529
; Sequence 529, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887

```

```

; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 529
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-529

Query Match      14.2%; Score 232.6; DB 17; Length 2106;
Best Local Similarity 56.6%; Pred. No. 2.6e-63;
Matches 504; Conservative 0; Mismatches 339; Indels 48; Gaps 2;

QY 170 GATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCACCCGCTGAACGTGACCTGCAACA 229
Db 273 GTTATGATCCAGGATAAGACCAAACTTCAAGGGCAATTCCTGTGTAGTAGTAGTCAACA 332
QY 230 TCTTCATCAACAGTTTTCAGCTCCGTCCCAAGACCAACAAATGGACTACCGGTGAATGTCT 289
Db 333 TTTTATTAACAGTTTTCAGTCCATTTCAAGAAACAACAAATGGACTATAGAGTTAAATCT 392
QY 290 TCTTGGGCAACAGTGGATGACCCAGCGCTCTCTTACCGAGAAATAT-----CCTGATG 343
Db 393 TCTGAGACAAAATGGAATGACCCAGCGCTGAAGCTCCCAAGTATTTTATGGAACCTGATTTT 512
QY 344 ACTCTCTGACCTCGATCCCTCATCTGCTGAGCTCTATCTTGAAGCCAGACCTCTCTTTTG 403
Db 453 ATGCACTGACAGTGGATCCCAATTTTCATGATGTGACCAAGTGTATGGAACCTGATTTT 572
QY 404 CTAAATGAGAAAGGGGCAACTTCCATGAGTGACCAAGCAACAAAGTACTTACGCGCATCT 463
Db 513 CAATGAAAGAAAGTGCCTTTTTCATGATGTGACCCAGGAAACATCTCTCTTTATTT 572
QY 464 TCAAGATGGGAATGTGCTGACAGCTGACCGCTGACCCCTCATTTTGTCTGCTGATGG 523
Db 573 TTCGATGAGAGATGCTTGTGAGATGAGGTATCTATTACTTTTCATGCTTTCATGCTTGG 632
QY 524 ACCTCAAGAACTTCCCATGGACATCCAGACGTGACCGATGACAGTGTGAGAGCTCATCCA 583
Db 633 ACTTGACATGTTTTCCTCATGGATACAAACGTTGCAAGATGCAACTGGAGAGCT----- 686
QY 584 TACTCTGCAGCCCTCTGCCATCTCTGTCTACTTTTCAGTTGGCTACACCATGAAGACCTCG 643
Db 687 -----TTGGTTTACAACTGATTTAC 710
QY 644 TGTTTGGTGGTGAAGATGCTCTCTGCTGTCCAAAGTGGCTGAGGGGCTGACTCTGCCCC 703
Db 711 GATTTATCTGGCAGTCAAGAGATCTCTGCTCAATAGAAAATTTGCTTGCCTCAATTTG 770
QY 704 AGTTTATCTTGGGGATGAGAAGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGA 763
Db 771 ATATCAAAAGGAAGATATTGAATATGTTAACTGTACAAAATATATAAGGCACGGCT 830
QY 764 AATTACCTGCTGAGGTAAAGTTTCACTTGAACGGCAGATGGGCTACTATCTGATTC 823
Db 831 ACTACACATGCTGGAAAGTCTCTTCCACCTGAGGAGGAGGTCGGCTTTTACATGATG 890
QY 824 AGATGCTGCTGCCCTCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 883
Db 891 GGGTCTACGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
QY 884 TGGATGCTGCCCTCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 943
Db 951 CGGACGCGAGTGTGCGCAGAGTGCCTTGGTATCTTCTCAGTCTCAGTCTGCTGCTGCT 1010
QY 944 AGAGCTCTGGCTCCCGGGCTCTTTGGCTTAAAGTGTCTTGAAGGCAATCGACATCT 1003
Db 1011 AGTGCAAACTTGGCTGAGCTTCCCAAGTTTCTCTATGTGAAGGCTCTTGTATGTTT 1070

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 03:05:59 ; Search time 6689.22 Seconds
(without alignments)

9332.242 Million cell updates/sec

Title: US-10-075-846-3

Perfect score: 1640

Sequence: 1 agacaactctgttcctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	45.1	1551	3	AK085727 Mus muscu
2	656	40.0	1294	9	AY418494 Homo sapi
3	654.6	39.9	1294	9	AY418496 Mus muscu
4	640.6	39.1	2387	3	AK078258 Mus muscu
5	638.2	38.9	3220	3	BC045547 Homo sapi
6	630.8	38.5	1294	9	AY418495 Homo sapi
7	561.8	34.3	1080	9	AY398900 Mus muscu
8	557.2	34.0	1080	9	AY398898 Homo sapi
9	477	29.1	1398	9	AY419225 Mus muscu
10	467	28.5	1398	9	AY419223 Homo sapi
11	458.8	28.0	1382	9	AY419224 Pan trogl
12	442.4	27.0	773	6	CD354404 UI-M-GMO-
13	401.8	24.5	662	1	AU169868
14	397	24.2	805	6	CD353729 UI-M-GMO-
15	378.6	23.1	540	7	CN266978 170004706
16	364.2	22.2	786	7	CO395810 AGENCOURT
17	346	21.1	760	7	CO421128 GSEZHC102
18	336.8	20.5	622	6	CD354268 UI-M-GMO-
19	311.4	19.0	821	7	CK354120 AGENCOURT
20	302.4	18.4	817	4	BG404477 60240604
21	295.6	18.0	678	1	AV729257 AV729257
22	292.4	17.8	779	5	BQ43713 UI-M-EMO-
23	287.2	17.5	715	9	AY398899 Pan trogl
24	274	16.7	624	7	CV024684 2154 Full

25 270 16.5 397 7 CN411646
26 266.6 16.3 992 2 BE981841
27 255.2 15.6 720 2 BE981459
28 239.6 14.6 552 4 BM440664
29 219.4 13.4 2875 3 AK083251
30 219.4 13.4 2991 3 BC027094
31 219.4 13.4 3008 3 BC037605
32 219.4 13.4 2593 3 BC022502
33 215.6 13.1 536 2 BF906462
34 211 12.9 3586 3 AK038618
35 209.4 12.8 2946 3 AK082578
36 203.8 12.4 895 6 CD301151
37 197.6 12.0 762 4 BI736322
38 193.8 11.8 1398 9 AY416040
39 190.6 11.6 1398 9 AY416042
40 185 11.3 1398 9 AY416041
41 178.2 10.9 1422 9 AY416045
42 177.4 10.8 1398 6 CD013911
43 176.2 10.7 1422 9 AY416043
44 175.6 10.7 1422 9 AY416044
45 175.4 10.7 937 7 CF408510

ALIGNMENTS

RESULT 1

AK085727

LOCUS

DEFINITION

AK085727

1551 bp mRNA linear HTC 03-APR-2004

Mus musculus 10 days lactation, adult female mammary gland cDNA,

RIKEN full-length enriched library, clone:D730029F11

product:glycine receptor, alpha 4 subunit, full insert sequence.

AK085727

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P. and Hayashizaki, Y.

AUTHORS

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalisation and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

11042159

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

REFERENCE

4

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

REFERENCE

5

The FANTOM Consortium and the RIKEN Genome Exploration Research


```

REFERENCE
AUTHORS
Mammalia: Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1294)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
JOURNAL
PUBMED
2 (bases 1 to 1294)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..1294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1294
/genes="GLR1"
/locus_tag="HCM6576"
ORIGIN
Query Match 40.0%; Score 656; DB 9; Length 1294;
Best Local Similarity 71.9%; Pred. No. 8.3e-172;
Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
QY 89 AGGAAGCTCAAAATCTGGAACCAAGGGGTCCAGCCCATGTCCCTCTGATTTCTCTAGACA 148
DB 15 AGGAGCTGNAGTGTCTGCTCGCACCCCAAGCTATGTACACCTCGGATTTCTTGATA 74
QY 149 AACTTATGGGGGAAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC 208
DB 75 AGCTAATGGGGAGAACCTCCGATATATGATGCCAGGATCAGGCCCAATTTTAAAGTCC 134
QY 209 CGGTGAACGTGACCTGCACATCTTATCAACAGTTTCAGTCCGTCCGTCACCMAGACCACAA 268
DB 135 CAGTGAACGTGACCTGCACATTTTATCAACAGTCTTGGTTCATTCGTGAGACAACCA 194
QY 269 TGGACTACCGGGTGAATGCTCTTGGCGCAACAGTGAATGACCAAGCTCTGCTTACC 328
DB 195 TGGACTATAGGGTCAACATCTTCTGGCGGAGCAATGGAACGACCCCGCTGGCTTATA 254
QY 329 GAGAATATCTGTGATGACTCTCTGGACCTCGATCCCTCCATGTGGACTCTATCTGGAAGC 388
DB 255 ATGAATACCTCTGACGACTCTCTGACCTGGACCCATCCATCTGAGTCCATCTGGAAC 314
QY 389 CAGACCTCTTCTTGTATGAGAAAGGGGCCAACTTCCATGAGGTGACACACGGAACAA 448
DB 315 CTGACCTGTCTTGTGCAACAGAGAAGGGGCCCACTTCCATGAGATCACCACAGACAAC 374
QY 449 AGTTACTGGGCATCTTCAAGAAAGGGAATGCTGTGTACAGCATCAGCTGACCTCATTT 508
DB 375 AATTGCTAAGGAATCTCCGGAATGGGAATGTCTCTACAGCATCAGATCACCCTGACAC 434
QY 509 TGTCTGCTGTGATGGACCTCAAGAACTTCCCCATGGACATCCAGAGTGCACGATCAGC 568
DB 435 TGGCTGCCCCATGGACTTGAAGAAATTTCCCATGGATGTCAGACATGATATCAAC 494
QY 569 TTGAGAGCTCATCCATCTCTGCGACCTCTGCGATCTCTGTCACTTTTCAGTTGGCTACA 628
DB 495 TGGAAAGCT-----TTGGATATA 512
QY 629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTGCTGCTCCAGTGGCTCAGG 688
DB 513 CGATGAATGACCTCATCTTTTGGTGGCAGGAACAGG---GAGCCGTGCAGGTAGCAGATG 569

```

```

QY 689 GSCTGACTGTGCCCCAGTTTATCTTCCGGATGAGAGGATCTAGGCTGTGTGTACCAAGC 748
DB 570 GACTAACTCTGCCCGCTTTATCTTGAAGGAAGAGAGGACTTGAGATACTTGACCACAGC 629
QY 749 ACTACAACACACGGGAAATTCACCTGCATCGAGTAAAGTTTACCTTGGAAACGGCAGATGG 808
DB 630 ACTACACACACAGTAAATTCACCTGCATGAGGCCCGGTTCCACCTGGAGCGGCAGATGG 689
QY 809 GTTACTATCTGATTCAGATGATCATCCCGAGCTACTCATCGTCACTCTCTGCTGCTCT 868
DB 690 GTTACTATCTGATTCAGATGATATTTCCAGGCTGTCTCATTTGTCTCTCTCATGATCT 749
QY 869 CTTCTGATCAACATGATGCTGCTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
DB 750 CTTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
QY 929 TCACCATGACACACAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
DB 810 TCACCATGACACACAGGAGCTCGGCTCTGAGCATCTCTGCCCAAGGTGCTCTATGTA 869
QY 989 AGCAATCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
DB 870 AAGCCATTGACATTTGGATGGCAGTTTGCCTGCTCTTGTGTTCTCAGGCCCTATTAGA 929
QY 1049 ATGCTGCCATAAATTTTGTCTGCTCAGCATAAAGAAATTCATACACTTCGAGAGGCG 1108
DB 930 ATGCTGCCATTAACTTTGTGCTCGGCAACATAAGGAGCTGCTCCGATTTCAGGAGGAG 989
QY 1109 AGAGGCGCCAAACGCTTTGGAGGAAGATATCATCCAGAAAGTCTGTTTCTATTTCCTGG 1168
DB 990 GGAGACATCACAGGAGGATGAAGCT-----GGAGAGGCCGCTTAACTTCTCTGCT 1043
QY 1169 ATGGCTTTGGGCC---ACTGCTTCAGGCAAGAGATGAGGTCCAATGGAAGGTTTC---TG 1222
DB 1044 ATGGGATGGGCCAGGCTGTCTACAGGCCAAGGATGSCATCTCAGTCAAGGGCGCAACA 1103
QY 1223 GATTTATAGTCCCAACCTCCAGGCCCTCTTCTAAGGAGAGGAGAAACACCGGGAAC 1282
DB 1104 ACAGTAAACACCAACACCCCTCTCCACCATCTAAGTCCCCCAGAGGAGATCGGAAAC 1163
QY 1283 TCTAGCTGACTGAGCCAAAGAGAAATTCACCAATCTCCCGGCTGCTTCCCTTTCACTT 1342
DB 1164 TCTTATCAGAGGGCCAAAGATCGACAATATCCGCGATTGGCTTCCCATGGCT 1223
QY 1343 TCCTCATCTTCAATATCTTCTACTGCGGTGTCTATAAAGTGTCTATGGTCAGAGATATCC 1402
DB 1224 TCCTCATTTCAACATGTTCTACTGATCATCTACAAGATTGTCCTGAGAGAGCGTCC 1283
QY 1403 ACCA 1406
DB 1284 ACAA 1287

```

```

RESULT 3
AY418496
LOCUS
DEFINITION
Mus musculus GLR1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY418496
VERSION
AY418496.1 GI:39774456
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1294)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
TITLE

```


Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillarary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2387)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saito, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
source
Location/Qualifiers
1..2387
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6J"
/db_xref="PANTOM,DB:6430545K04"
/db_xref="taxon:10090"
/clone="6430545K04"
/sex="male"
/tissue type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
466..1813
misc_feature
/notes="glycine receptor, alpha 1 subunit (MGD|MGF:95747, GB|NM_020492, evidence: BLASTN, 99%, match=1348) putative"
ORIGIN
Query Match 39.1%; Score 640.6; DB 3; Length 2387;
Best Local Similarity 72.4%; Pred. No. 2.1e-167;
Matches 939; Conservative 0; Mismatches 299; Indels 59; Gaps 6;
QY 116 CCAGGCCATGTCCTCGATTCCTAGACAACTTATAGGCGCAACATCTGGATATG 175
|||||
|||||

563 CCAGCCTATGTCAACCTCGGACTTCTCTGGATAGCTCATGGGAGGACTTCTGGGTATG 622
QY 176 ATGCCAGGATTCGGCCCAATTTTAAAGCCCAACCGGTGAACGAGCTGCAACATCTTCA 235
DB 623 AGCCAGGATCAGGCCCAATTTTAAAGTCTCTCTGTGAATGTAAAGTTGCAACATCTTCA 682
QY 236 TCAACAGTTTACGTCCTCGTCAACAGACCAACATGAGCTACCGGGTGAATGTTCTTTC 295
DB 683 TCAACAGTTTTCGTTTCCATCGCTGAGACCAACATGAGCTATAGGGTCAACATCTTCTCTGA 742
QY 296 GGCACACAGTGGGAATGACCCACGCTGTCTCTACCGAGAAATATCTCTGATGACTCTCTGGACC 355
DB 743 GGCAGCAGTGGGAACGACCCCGTCTGGCTCAATGAATACCTGATGACTCTCTGGACC 802
QY 356 TCGATCCCTCCATGTGGACTCTATCTGGAAGCCAGACCTCTTCTTTTGTAAATGAGAAAG 415
DB 803 TTGACCCATCTATGTTGGATTCCATCTGGAAGCTGACTTGTCTTTGCAATGAGAAGG 862
QY 416 GGGCCCAACTTCCATGAGGTGACCGGACCAACAGTTACTGCGCATCTTTCAGAGATGGGA 475
DB 863 GGGCCCACTTCC--GAAATCACCGGACCAACAAACTGCTAAAGAACTCTCCCGGAATGGCA 920
QY 476 ATGTGCTGTACAGCATCAGCTGACCCCTCATTTTGTCTCTGCTGATGGACCTCAAGAACT 535
DB 921 ATGTCTCTACAGCATCAGAAATCAACCTGACGCTGGCTGCCCCATGGACCTGAAGAAAT 980
QY 536 TCCCATGACATCCAGACGTCGACGATGACGCTGAGAGCTCATCCATCTCTGCAGCC 595
DB 981 TCCCGATGATGTACAGACGCTGATCATGCAACTCGAAAGCT----- 1022
QY 596 CTCTGCCATCTCTGTCACTTTTCAGTTGGCTACACCATGAAGACCTCTGTTGAGTGGC 655
DB 1023 -----TTGGATATACCATGAAGACCTCATCTTTGAGTGGC 1058
QY 656 TGAAGATGCTCTCTGTCTCAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGC 715
DB 1059 AGAGCAAG---GAGCTGTGAGGTGGCAGACGACTGACCTGCTCTGATTTATTTCTGA 1115
QY 716 GGGATGAGAAGGATCTAGGCTGTGTGTACCAAGCCTTACAAACAGGAAATTCACCTGCA 775
DB 1116 AGGAAGAGAAGACCTGAGATATCTGCACCAAGCCTTACAAACAGGTAATTCACCTGCA 1175
QY 776 TCGAGTAAAGTTTACCTGGAAACGCGAGATGGGCTACTCTGATTCAGATGTATATCC 835
DB 1176 TCGAGGCCGATTTCCACCTGGAGCGCAGATGGGCTACTACCTGATTCAGATGTATATCC 1235
QY 836 CCAGCCTACTCATCTGTCATCTCTGCTGCTGCTCTCTCTTGGATCAACATGGATGCTGCC 895
DB 1236 CCAGCCTGCTCATCTGTCATCTCTGCTGCTGCTCTCTTCTGGATCAACATGGATGCTGCAC 1295
QY 896 CTGCCCGTGTGGGCTGGGCTATCACACCGTGTCTCAACATGACCAACCCAGAGCTCTGGCT 955
DB 1296 CAGCTCGTGTGGGCTGGGCTATCACACAGTGTCTCAACATGACCAACACAGAGCTCTGGCT 1355
QY 956 CCGGGCTCTTTTGGCTTAAGGTGTCTAGTGAAGCAATCGACATCTCGATGGCTGTGT 1015
DB 1356 CCGGAGCTCTCCCTACCCAGGTGTCTACGTGAAAGCTATTGACATTTGGATGGCTGTTT 1415
QY 1016 GTCGCTCTTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075
DB 1416 GCTGCTCTTCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1475
QY 1076 AGCATAAAGAAATTCATACGACTTTCGAAGAGGCGAGCGGCCCAACCTTTGGAGGAAGATA 1135
DB 1476 AACACAGGA-----ACTGCTTCGATTTAGGAGGAAGGCGACATCAAGGATGATG 1529
QY 1136 TCATCAAGAAAGTGTGTTTCTATTTCCGTGGCTATGCTGCTGGGCC---ACTGCTCGAGG 1192
DB 1530 AGGGTGGAGAGGACCTTCAACTTCTCTGCTTATGGGATGGGCCCGCAGCTGTCTCGAG 1589
QY 1193 CAAGAGATGGAGGTCCATGGAAGTTCTGGGATTTATAGTCCC---CAACCTCCAGCCC 1249
DB 1590 CCAAGGATGGCATCTCTGTCAAGGGTGCACCAACAAACCAACCACTAAACCCGCTCTCTG 1649

QY 1250 CTCTTCTAAGGAAGGAGAAACACCGGAAACTCTACGTGGACTGACCAAGAGAAATTG 1309
 Db 1650 CGCCATCAAGTCCCGGAGGAGATCGGAAACTCTTCCAGAGAGCAAGAAGATCG 1709
 QY 1310 ACACATCTCCCGGGTGTCTCCCTTTTCATTTCTCTCTCAATCTCTCTACTGGG 1369
 Db 1710 ACAAGATATCTCGCATCGGTTTCCCATGCGCTTCTCTCTCTCAATCTCTCTACTGGA 1769
 QY 1370 TTGTCTATAAAGTGCTATGCTCAGAGATATCCACCA 1406
 Db 1770 TCATCTATAAGTCTGCTCCGAGAGAGATGTCCACAA 1806

RESULT 5
 BC045547 3220 bp mRNA linear HTC 12-OCT-2004
 LOCUS Homo sapiens cDNA clone IMAGE:5302075, containing frame-shift errors.
 DEFINITION BC045547
 ACCESSION BC045547.1 GI:28374243
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 3220)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.E., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 3220)
 DIRECTOR MGC Project.
 DIRECT SUBMISSION
 SUBMITTED (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT: MGC help desk
 EMAIL: cgaps-remail.nih.gov
 TISSUE Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 CONTACT: amadnan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 106 Row: p Column: 9
 This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 40254460
 This clone has the following problem: frame shifted.
 FEATURES
 Location/Qualifiers
 1..3220
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5302075"
 /tissue_type="Brain, hypothalamus"
 /clone_lib="NIH MGC_96"
 /lab_host="DH105"
 /note="Vector: pBluescript"

ORIGIN
 Query Match 38.9%; Score 638.2; DB 3; Length 3220;
 Best Local Similarity 70.0%; Pred. No. 1.1e-166;
 Matches 942; Conservative 0; Mismatches 348; Indels 55; Gaps 4;
 QY 63 CCTCTCAGGCTGGCTTGGCAAAAGAGAAAGTCAAATCTGGAACCAAGGGTCCAGCC 122
 Db 600 CTTGAGGCGGCTTTCTGCAAAAGACCATGACTCCAGGTCTGGAACCAACACCTTTCAGAGC 659
 QY 123 CATGTCCCTCTGATTTCTAGACAAACTTATGGGCGAAACATCTGGATATGATGCCAG 182
 Db 660 CCTATCTCTTCAGATTTCTTGGCAAGTTAATGGGAAGGACATCAGGATATGATGCAAG 719
 QY 183 GATTGGGCGCAATTTTAAAGGCGCCACCCCTGAAGCTGACCTGCAACATCTTCAATCAACAG 242
 Db 720 AATGAGCCCAATTTTAAAGGCTCTCCAGTAAACGTTACTTGCATATATTTTATCAACAG 779
 QY 243 TTTTCAAGTCCGTCACCAAGACCAATGACCTACCGGGTGAATGTCTTCTGGCGCAACA 302
 Db 780 TTTTGGATCAGTCACAGAAACGACATGACCTACCGAGTGAATATTTTCTGAGACAACA 839
 QY 303 GTGGAATGACCAAGCTGTCTTACCGAGAAATCTCTGATGACTCTCTGAGACCTCGATCC 362
 Db 840 GTGGAATGATTCACGGCTGGCGTACAGTACGATGACCCAGATGACTCCCTGGACTTGGACC 899
 QY 363 CTCCATGCTGGACTCTATCTGGAAGCGACAGCTCTCTTTGTCTAATGAGAAAGGGGCCAA 422
 Db 900 ATCCATGCTAGACTCCATTTGGAAACAGATTTGTTCTTTGGCAATGAGAGGGTCCCAA 959
 QY 423 CTTCCATGAGGTGACCAAGCAACCAAGTTACTGGCATCTTCAAGAAATGGGAATGTGCT 482
 Db 960 CTTCCACGATGATCACCACCTGACAAACAAATGCTACGGAATTCGAAAAATGGCAAGTGT 1019
 QY 483 GTACAGCATCAGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCAT 542
 Db 1020 CTACAGTATCAGACTCACCTTGCACCTTATCTCTGCGATGAGACTTGAAGAACTTCCGAT 1079
 QY 543 GGACATCCAGAGCGTCACGATGACGCTTCAGAGCTCATCCATCTCTGAGCCCTCTGCC 602
 Db 1080 GGATGTCAGACTGTACAACTGACGCTGGAGAGTT----- 1114
 QY 603 ATCTCTGTCACTTTTCAGTTGGCTTACACCAAGAAAGCTCGTGTGTTGAGTGGCTGGAAGA 662
 Db 1115 -----TTGGGTACACGATGAATGACCTGATATTTGAGTGGTAAAGTGA 1157
 QY 663 TGCTCTGCTGTCGAAGTGGCTGAGGGGCTGACTCTGCCCGCAGTTTATCTTGGGGATGA 722
 Db 1158 TGGTC---CAGTGCAGAGTTGCTGGAAGGATGACCCCTGCCCGCAGTTTATTTTGAAGAAGA 1214
 QY 723 GAAGGATCTAGGCTGTGTGACCAAGCACTACAAACACAGGGAATTCACCTGATCGAGGT 782
 Db 1215 GAAGGACTTGGCTACTGTACAAAGCACTACAACTGGAAGTTTACCTGATTCAGGT 1274
 QY 783 AAAGTTTCACTGGAACCGCAAGATGGGCTATCTATCTGATTCAGATGTATATCCCGACCT 842
 Db 1275 CAAGTTTCACTGGAACCGCAAGATGGGATATTTTGTATCCAGATGTACAT-CCGACCT 1333
 QY 843 ACTCATGCTATCTGCTGCTGGCTCTCTTCTGGAATCAACATGATGCTGCCCTCCCG 902
 Db 1334 GCTTATAGTAATTTTGTCTCTGGGTTTCTTTTGGATAAATATGATGACAGCCCTCCAG 1393

```

QY 903 TGTGGCCCTGGGATCACCACCGTGTCTACCATGACCACCGAGACTCTGGCTCCGGGC 962
Db 1394 GGTGCGCATGGGATCACCACAGTCTTAACGATGACCACCGGAGTTAGGCTCCAGGC 1453
QY 963 CTCCTTTGCTAGGCTCTAGTGAAGCAATCGACATCTGATGGCTGTGTCTGCT 1022
Db 1454 ATCTTGCCAAAGTCTCTATGTAAGGATGACATCTGATGGCGGTGCTCTCT 1513
QY 1023 CTTTGTGTTGCTGCTTCTGCGAGTATGTCGCATAAAATTTTGTCTCGTCAGCATAA 1082
Db 1514 GTTTGTGTTGCTGCTTACTGGAATACGAGCGGTGAATCTGCTCCAGGCAACACAA 1573
QY 1083 AGAATTCATAGACTTCGAAGAGGAGGAGGCGCCCAACGCTTGGAGGAAGATATCATCA 1142
Db 1574 GGAAGTTCCTGCGCTCCGGAAGAGAGAGAGGAGGAGAAATAGGAAGAGAGCGTTACTCG 1633
QY 1143 AGAAGTCTGTTCTATTTCTGCTGCTATGCTTGGCGCACTGCTGCGCAAGAGATG 1202
Db 1634 TGAAGTCTGTTTAAATTTTACGCTTATGGATGGGTCACTGCTCCAAAGTGAAGATGG 1693
QY 1203 AGTCCAAATGGAGGTTCTGGCATTTATATGTCCTCCCAACCTCCAGCCCTCTTTCTAAGGA 1262
Db 1694 AACAGCTGTCAAGGCCAC-----ACCTGCCAACCCCACTCCCAACCGCCAAAGA 1744
QY 1263 AGAGAAACACCGCGAACTCTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1322
Db 1745 TGAGATGCTATCAAGAAGAGTTTGTGGACCGGCAAGAGGATTTGACACGATATCTCG 1804
QY 1323 GGTGCTTCTCCCTTCACTTCTCTCATCTTCAATATCTTCTACTGGTGTCTATAAGT 1382
Db 1805 AGTGGCTTCCCATGCGCTTCTCTCATTTTCAACATCTTTTACTGGATCACATACAGAT 1864
QY 1383 GCTATGCTCAGAGATATCCACAG 1407
Db 1865 CATTCGGCATGAGATGTCACAG 1889

RESULT 6
AY418495 1294 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes GLRA1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY418495
VERSION AY418495.1 GI:39774455
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
2 (bases 1 to 1294)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1294
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"

```

```

/db_xref="taxon:9598"
<1..>1294
/genes="GLRA1"
/locus_tag="HCM6576"

gene
ORIGIN

Query Match 38.5%; Score 630.8; DB 9; Length 1294;
Best Local Similarity 67.2%; Pred. No. 9.4e-165;
Matches 886; Conservative 0; Mismatches 387; Indels 45; Gaps 2;

QY 89 AGGAAGTCAAAATCTTGAACCAAGGGGTCCAGCCCATGTCCCGCTCTGATTTCTAGACA 148
Db 15 AGAGGCTGAAGTGTCTCGCTCCGACCAAGCCTATGTACCCCTCGGATTTCTTGATA 74
QY 149 AACTTATGGGGGAACATCTGATATGATGCGAGGATCGGCCCAATTTTAAAGGCCAC 208
Db 75 AGCTAATGGGAGAACCTCCGATATGATGCGAGGATCAGGCCCAATTTNNNNNGTCCCC 134
QY 209 CCGTGAACGTGACCTGCAACATCTTCAACAGTTTCAAGTCCGTCAGCCGTCACCAAGACCACAA 268
Db 135 CAGTGAACGTGAGCTGCAACATTTTCAACAGCTTTTGGTTCATTTGTGTGAGAACCA 194
QY 269 TGGACTTACCGGTGAATGTCTTCTGCGGCAACAGTGGAAATGACCCACGCTCTCTTACC 328
Db 195 TGNNNNNNGGTCAACNNCTTCTGCGGAGCAATGGAACGACCCCGCTGCGCTATA 254
QY 329 GAGAATATCTGATGATCTCTGGACTCGATCCCTCCATGTGTGGAATCTTATCTGGAAGC 388
Db 255 ATGAATACCTCTGATGATCTCTGGACTCGATCCCTCCATGTGTGGAATCTTATCTGGAAC 314
QY 389 CAGACTCTCTTCTGTAATGAGAAAGGGGCAACTTCCATGAGGTGACCAAGGCAACA 448
Db 315 CTGACCTGTCTTCTGCAAGAGAGGGGGCCCACTTCCATGAGATGACCAAGAACAA 374
QY 449 AGTTACTGCGCATCTTCAAGAAATGGGAATGTGTGTGACAGCATGAGCTGACCTCATTT 508
Db 375 AATTGCTAGGATCTCCCGAATGGGAATGTCTCTAGAGCATCAGATCACTCCCTGACAC 434
QY 509 TGTCTGCTGATGAGCACTCAAGAACTTCCCATGAGATGACAGAGCGTGACGATGACG 568
Db 435 TGGCTGCCCCATGAGCTTGAAGAAATTTCCCATGAGATGTCACAGACATGTATCATGCAAC 494
QY 569 TTGAGAGCTATCCATCTCTGCGACCTCTGCCATCTCTGTACATTTTCAAGTTGGCTACA 628
Db 495 TGGAAAGCT-----TTGGATATA 512
QY 629 CCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTCTGTCTCCAAAGTGGCTGAGG 688
Db 513 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCGGTGACAGTACAGATG 569
QY 689 GGCTGATCTGCCCCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTGTGACCAAGC 748
Db 570 GACTAACNNNGCCAGTTTANNNNNAAGGAAGAGAGGACTTTGAGATNNNTGACCAAGC 629
QY 749 ACTACAAACAGGGAATTCACCTGATCGAGTAAAGTTTACCTTGGACCGCAGATGG 808
Db 630 ACTACAAACAGGTAATTCACCTGATTTGAGGCCCGGTTCCACCTTGGAGCGCAGATGG 689
QY 809 GCTACTATCTGATTCAGATGTATCATCCAGCCCTACTCATCGTCTCTCTCTCTGGGTCT 868
Db 690 GTTACTACTGATTCAGATGTATATTCACGACCTGTCTCATTTGTCATCTCTCATGATCT 749
QY 869 CTTCTGGATCAACATGATGTCGCCCTTCCCGGCTCTTTGGCTTAGGTGTCCTAGCTGA 988
Db 750 CTTCTGGATCAACATGATGTCGCCCTTCCCGGCTCTTTGGCTTAGGTGTCCTAGCTGA 809
QY 929 TCACCATGACCAACAGATGTCGCCCTTCCCGGCTCTTTGGCTTAGGTGTCCTAGCTGA 988
Db 810 TCACCATGACCAACAGATGTCGCCCTTCCCGGCTCTTTGGCTTAGGTGTCCTAGCTGA 869
QY 989 AGGCAATGACATCTGATGGCTGTGTCTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1048
Db 870 AGGCAATGACATCTGATGGCTGTGTCTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 929

```

QY	1049	ATGCTGCCATAAATTTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAAGGC	1108	Db	84	AGACCATGACTCCAGGTCTGGAAACATCCCTCGCAGACCCCTATCTCCTTCAGATTTCTT	143						
Db	930	ATGCTGCCGTAACTTTGTTCTCGCAACATAAGGAGTCTCCGATTCAGGAGAAGC	989	QY	144	AGACAACTTATGCGCGCAACATCTGGATATGATGCGAGGATTCGCGCCCAATTTTAAAGG	203						
QY	1109	AGAGGCGCAACGCTTGGAGGAAGATATCATCCAGAAAGTCGTTCTTATTTCCGTGGCT	1168	Db	144	GGATAAATTAATGGGAAGAACATCAGGATATGATGCAAGAAATCAGGCCAATTTTAAAGG	203						
Db	990	GGAGACATCACAGNN	1049	QY	204	CCACCCGCTGAACGTGACCTGCAACATCTTCATCAACAGTTTTCAGTCCGTCACCAAGAC	263						
QY	1169	ATGGCTTGGCCACTCCCTCGAGGCAAGAGATGGAGTCCAATGGAAAGTTTCTGGCATTT	1228	Db	204	TCCTCCAGTAAACGTTTACTTGGCAATATTTTATCAACAGTTTGGATCAGTCACAGAAAC	263						
Db	1050	TGGGCCACAGCTGTCTACAGGCAAGGATGGNNNNNNNNNNNNNNNNNNNNNNNN	1109	QY	264	CACATGAGTACACCGGTGAATGTCTCTTGGCGCAACAGTGAATGATCCAGCGCTGTC	323						
QY	1229	ATAGTCCCCCAACCTCCAGGCGCTCTTCTAAGGAAGGAGAAACACCGCGAACTCTACG	1288	Db	264	CACCATGAGTACACCGGTGAACATTTTCTGAGACACAGTGGAAATGATTCACGCGCTGGC	323						
Db	1110	ACACCAACCAACCCCTCTCTGACCATCTAAGTCCCCAGAGGAGTGCAGAAATCNCNA	1169	QY	324	CTACCGAGAATATCTGATGACTCTCTGACCTCGATCCCTCATGCTGAGTCTCTATCTG	383						
QY	1289	TGAGCTGAGCAGAGAAATGACACATCTCCGGGCTGTCTTCCCTTTTACCTTCTCTCA	1348	Db	324	ATACAGTGAATACCCAGATGATTCCTGATTTGGATCCCTCAATTTGGATTCGATTTG	383						
Db	1170	TCAGAGGCGCAAGAGATCGACAAATATCCGCAATTCGCTTCCCATGGCTTCCCTCA	1229	QY	384	GAAGCCAGACCTCTTCTTTTGTCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCACGA	443						
QY	1349	TCTTCAATATCTTCTACTGGGTGTCTATAAGTGTCTATAGTGTCTGTCAGAGATATCCACCA	1406	Db	384	GAACCGGATTTGTTCTTTGCCAATGAAAGGGTGCCAAATTTCCATGATGTCCACCTGA	443						
Db	1230	TTTTCACATGTTCTACTGGATCATCTACAAGATTGTCGTAGAGAGGACGTCACAA	1287	QY	444	CAACAGTTTATGCGGATCTTCAAGAAATGGGAATGTGCTGTACAGATCAGCTGACCT	503						
RESULT 7				Db	444	CAACAGTTTATGCGGATTTCCAAAATGGCAAGTGTCTCTACAGTATTAGACTCACCTT	503						
LOCUS	AY398900	1080 bp	DNA	linear	GSS 15-DEC-2003								
DEFINITION	Mus musculus GURA2 gene, VIRTUAL TRANSCRIPT, partial sequence,												
ACCESSION	AY398900												
VERSION	AY398900.1	GI:39754889											
KEYWORDS	GSS.												
SOURCE	Mus musculus (house mouse)												
ORGANISM	Mus musculus												
REFERENCE	1 (bases 1 to 1080)												
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Slnsky, J.J., Adams, M.D. and Cargill, M.												
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios												
JOURNAL	Science 302 (5652), 1960-1963 (2003)												
PUBMED	14671302												
REFERENCE	2 (bases 1 to 1080)												
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Slnsky, J.J., Adams, M.D. and Cargill, M.												
TITLE	Direct Submission												
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA												
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.												
FEATURES	Location/Qualifiers												
source	1..1080												
	/organism="Mus musculus"												
	/mol_type="genomic DNA"												
	/db_xref="taxon:10090"												
	<1..>1080												
	/gene="GURA2"												
	/locus_tag="HCM0044"												
gene													
ORIGIN													
Query Match	34.3%; Score 561.8; DB 9; Length 1080;												
Best Local Similarity	73.7%; Pred. No. 1.8e-145;												
Matches	763; Conservative 0; Mismatches 227; Indels 45; Gaps 2;												
QY	84	AAAGAGGAGTCAATCTTGAACCAAGGGTCCAGGCCCATGTCCCTCTGATTTCTT	143	RESULT 8									
				AY398900									

LOCUS AY398898 1080 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens GLRA2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY398898
VERSION AY398898.1 GI:39754887
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1080)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1080
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>1080
/gene="GLRA2"
/locus_tag="HCM0044"
ORIGIN
Query Match 34.0%; Score 557.2; DB 9; Length 1080;
Best Local Similarity 72.7%; Pred. No. 3.6e-144;
Matches 768; Conservative 0; Mismatches 243; Indels 45; Gaps 2;
QY 63 CTTCTCAGGTCGGCTTGTGCAAAAGAGGAGTCAATCTGGAACCAAGGGGTCCAGCC 122
DB 63 CTTTCAGGACGGCTTCTGCAAGACCATGACTCCAGGTCTGGAACCAACACTTTCACAGAC 122
QY 123 CATGTCCTCTGATTTCTAGACAACTTATGGGGCGAACATCTGGATATGATGCCAG 182
DB 123 CCTATCTCTTCAGATTTCTTGACAAAGTTAATGGGAAGGACATCAGGATATGATCAAG 182
QY 183 GATTCCGGCCCAATTTTAAAGGCCACCGGTGAACGTGACCTGCAACATCTTTCATCAACAG 242
DB 183 AATCAGGCCCAATTTTAAAGTCTCTCCAGTAAAGTACTTGCATATTTTATCAACAG 242
QY 243 TTTTCAGTCCTGTCACCAAGACCAATATGACTACCGGGTGAATGTCTTCTTGGCGCAACA 302
DB 243 TTTTGGATCAGTCACAGAAACGACCATGGACTACCGAGTGAATATTTTCTGAGACAACA 302
QY 303 GTGGATGACCCAGCGCTGCTCCACGAGATATCTGTGATGATCTCTGGACCTTCATCC 362
DB 303 GTGGATGATTTACCGGTGGCGTACAGTGAAGTACCCAGATGACTCTCCGTGGACTTGGACCC 362
QY 363 CTCATGCTGAGTCTATCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGCCAA 422
DB 363 ATCATGCTAGATCTCATTTTGGAAACAGATTTGTTCTTGGCAATGAGAGGGTGCCTAA 422
QY 423 CTTTCCATGAGGTGACCAAGCAACCAAGTTATCTGCGCATCTTCAAGATGGGAATGTGCT 482
DB 423 CTTTCCAGATGTCAACCACTACCAACAAATTTGCTACCGATTTTCGAAAAATGGCAAGTCT 482
QY 483 GTACAGCATCAGCTGACCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCAT 542

483 CTACAGTATCAGACTCACCTTGACCTTATCTCTGCTCCATGGAGCTTGAAGAACTTTCCGAT 542
QY 543 GGACATCCAGACGTGACGATGCGAGCTTGAGAGCTCATCTGAGCCCTCTGCC 602
DB 543 GGATGTCACAGACCTGTACAAATGACGCTGGAGAGTT----- 577
QY 603 ATCTCTGTACATTTTCAGTTTGGCTTACACCATGAAGACCTCGTGTGTTAGTGGCTGAAGA 662
DB 578 -----TTGGGTACACGATGAATGACCTGTATATTTGAGTGGTAAAGTGA 620
QY 663 TGTCTCTGCTGTCACCAAGTGGCTGAGGGGTGACTCTGCCCCCAGTTTATCTTGGGGATGA 722
DB 621 TGGTC---CAGTGCAGATGCTGAGGAGTGTGACCTGCCCCAGTTTATTTTGAAGAAGA 677
QY 723 GAAGATCTAGGCTGTGTGACCAAGCACTACAAACAGGGAATTCACCTGCAATCAGGT 782
DB 678 GAAGAACTTGGCTACTGTACAAAGCACTACAAACACTGGAAGATTTTACCTGCAATGAGGT 737
QY 783 AAAGTTTACCTGGACGCGAGATGGCTACTATCTGATTCAGATTCAGATCCCCAGCCT 842
DB 738 CAAGTTTTCATCTGGAACGCGCAATGGGATATTATTTGATCCAGATGATATCCCAAGCCT 797
QY 843 ACTCATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
DB 798 GCTTATAGTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
QY 903 TGTGGGCTTGGGATCACCACCGTGTCTACCATAGCACCCAGAGCTCTGGCTCCCGGGC 962
DB 858 GGTGCGACTTGGGATCACCACAGTCTTAAAGCATGACACCCAGAGTTCAGGCTCCAGGC 917
QY 963 CTCCTTGGCTTAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
DB 918 ATCTCTGCGCAAGGCTCTCTATGTAAGGCGATGATGATGATGATGATGATGATGATGATGAT 977
QY 1023 CTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
DB 978 GTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037
QY 1083 AGAATTCATACGACTTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1118
DB 1038 CGAGTTTCTGCGCTCCGAAAG 1073
RESULT 9
AY419225 1398 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus GLRA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY419225
VERSION AY419225.1 GI:39775185
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1398)
AUTHORS Todd,M.A., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1398)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering

ORIGIN		/locus_tag="HCM6819"	
Query Match		28.5%; Score 467; DB 9; Length 1398;	
Best Local Similarity		61.3%; Pred. No. 7e-119;	
Matches		830; Conservative 0; Mismatches 453; Indels 72; Gaps 6;	
QY	80	TGCGAAGAGGAGTCAATCTGGACCAAGGGTCCAGCCCATGTCCTCTGATT	139
DB	77	TTCCCAAGAGGAAACAGACAGTGCAGATCTCGAAGTCTCAATGTACCTTCTGATT	136
QY	140	TCCTAGACAACTTATGGGCGCAACATCTGGATATGATGCCAGATTCGGCCCAATTTTA	199
DB	137	TTCTGATAATTAATGGGCGAGACATCAGATATGATGCAAGATCAAGCCCAATTTTA	196
QY	200	AAGGCCACCCGTAACGTCGACCTGCAACATCTTTCATCAACAGTTTCAGCTCCGTCA	259
DB	197	AAGNN	256
QY	260	AGACCAATAGGACTACCGGTGAATGTCTTCTGGGGCAACAGTGGATGACCCAGCC	319
DB	257	NN	316
QY	320	TGCTCTACCGAGATATCTCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGACTTA	379
DB	317	TCGGGTACGTGAATCTCTGACGACTCTTTAGACCTCGACCCCTCATGTTGGATCCA	376
QY	380	TTTGSAGCCAGACTCTCTTTGCTAATGAGAAAGGGCCAACTTCCATGAGGTGACCA	439
DB	377	TTTGGAAACCTGATTTGTTCTTTGGCCAAATGAAAGGGTCCCACTTTCATGAAGTCACTA	436
QY	440	CGGACAAAGATTACTGCGCATCTTCAAGATGGGAATGTGTGTACAGCATCAGG---C	496
DB	437	CAGACAACTAATGCTAAGAAATTTCAAAAATGGAATGTCTTTTATCAATAAGGAGAT	496
QY	497	TGACCTCATTTTGTCTGCTGATGACCTCGACCTCAAGAACTTCCCATGGACATCAGAGT	556
DB	497	TAACATTAACACTTCTGTCGAATGATCTCAAGAAATTTCCCATGGATGTCAAAACAT	556
QY	557	GCACGATGACGCTGAGAGCTATCCATCTCTGCGACCCCTCTGCCATCTCTGTCACTTT	616
DB	557	GTATAATGCAACTGGAAGCT-----	577
QY	617	CAGTTGGCTACACATGAAGAACTGTGTGTGAGTGGCTGGGAAGATGCTCTGTGTCTC	676
DB	578	---TTGGGTACAAATGAATGATCTCTATTTTGAATGCG---AAGATGAGGCACCCGTAC	631
QY	677	AAGTGGCTGAGGGCTGACTCTGCCAGTTTATCTTGGCGATGAGAAGATCTAGGCT	736
DB	632	AAGTGGCAGAGGACTCCTTTGCCAGTTTCTGTGTGAAGAAAGAAAGATTTAGCAT	691
QY	737	GTTGTACCAAGCACTACAAACAGGGAATTCACCTGCAATCGAGGTAAAGTTTCACTGG	796
DB	692	ACTGCACTAAACATTACATACAGGAAGTTTACGTGTATAGATGCGATTCATCTGG	751
QY	797	AAGCGAGATGGGCTACTATCTGATTCAGATGATACATCCAGCTACTCATGTCATCC	856
DB	752	AGCGCAAAATGGGATATCTATCTGATCCAGATGATACATTCCTGCTCTCTGATTTATTC	811
QY	857	TGTCCTGGTCTCTCTGATCAACATGATGCTGCCCTGCCCTGCTGGGCTGGCA	916
DB	812	TATCTCTGGGTTCTGTTCTGATCAACATGATGATGATGATGATGATGATGATGATG	871
QY	917	TCACCACTGCTCACCATGACCAACAGAGCTCTGCTCCGGGCTCTTTCCTTAAGG	976
DB	872	TAACCACTGCTCAACATGATCTACAGATGTTTACAGATCAAGCTTCTTCCCAAG	931
QY	977	TGTCCTAAGGCAAGCAATCGATCTGGATGCTGTGTCTGTCTTTGTGTTGCTG	1036
DB	932	TTTCATATGTCAAAGCTATTGATATTGGATGTCAGTATGCTCTCTTTTGTGTTTCAG	991
QY	1037	CCTTGTGGATGCTGCTCAATTTTGTCTGTCAGCATGAAGATTCATAGGAC	1096
Db		992 CACTTCTGGATATGCAGCTGTAAATTTTGTATCAAGACACACAAAGAACTTCTGAGAT	
QY		1097 TTCGAAGAAGCAGAGGCGCCA-----AGCTTGGAGGAAGATATCATCC	
Db		1052 TTCGACGAAGAAGAAATAAGACAGAAAGCTTTTTCAGATGGGAAATTTTACCGTTCT	
QY		1142 AAGAAAGTCTTCTATTTTCGTTGGCTATGGCTTGGGCGCACTGCCTGCGAGGAAGATG	
Db		1112 CAGATATGATGATGAGTGAAGGAAAGCGGATTCAGCTTTCAGCCTATGGATGGAC	
QY		1202 GAGGTCCAATGGAAGTTTCTGGCAT-----TTATAGTCCCAACCTTCCAGCCCC	
Db		1172 CATGCTCTACAGCAAGAGGATGCGATGACTCCAAAGGCGCCCAACCTGTCCAGTAA	
QY		1253 TTCGAAGGAAGAGAAACACGCGGAAACTCTACGTGACTGAGCAAGAAATTTGACA	
Db		1232 TGCAAAAAGTCTCTGATGAATGAGGAAGTCTTTATCGACCGGGCAAGAAATGATA	
QY		1313 CCATCTCCCGGCTGTCTTCCCTTTCCATCTTCTCATCTTCAATATCTTCTACTGGTTG	
Db		1292 CCATCTCCCGGCTGTCTTCCCTTTCCATTTAGATTTTATATATTTTCTACTGGTTA	
QY		1373 TCTATAAGTGTCTATGCTCAGAAAGATATCCACCAG	
Db		1352 TCTATAAATTTCTAGGCATGAGGATATTCATCAG	
RESULT 11		AV419224 1382 bp DNA linear GSS 17-DEC-2003	
LOCUS		Pan troglodytes GLRA3 gene, VIRTUAL TRANSCRIPT, partial sequence,	
DEFINITION		genomic survey sequence.	
ACCESSION		AY419224	
VERSION		AY419224.1 GI:39775184	
KEYWORDS		GSS.	
SOURCE		Pan troglodytes (chimpanzee)	
ORGANISM		Pan troglodytes	
REFERENCE		1 (bases 1 to 1382)	
AUTHORS		Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.	
TITLE		Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios	
JOURNAL		Science 302 (5652), 1960-1963 (2003)	
PUBMED		14671302	
REFERENCE		2 (bases 1 to 1382)	
AUTHORS		Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.	
TITLE		Direct Submission	
JOURNAL		Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT		These sequences were made by sequencing genomic exons and ordering them based on alignment.	
FEATURES		Location/Qualifiers	
source		1..1382	
		/organism="Pan troglodytes"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9598"	
gene		<1..>1382	
		/gene="GLRA3"	
		/locus_tag="HCM6819"	
ORIGIN			
Query Match		28.0%; Score 458.8; DB 9; Length 1382;	
Best Local Similarity		61.0%; Pred. NO. 1.4e-116;	
Matches		824; Conservative 0; Mismatches 455; Indels 72; Gaps 7;	
QY		80 TGCAAAAGAGGAAGTCAATCTGGAAACCAAGGGTCCAGCCCATGTCCTCTGATT	

Db 77 TTGCCAAGGAAACAGATAGTGAAGATCTGAAAGTGTCTCAATGTCACTTCTGATT 136
Qy 140 TCTAGACAACTTATGGGCGCAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTA 199
Db 137 TCTGTGATAATTAATTTGGGCGAGACATCAGGATATGATCAAGAAATCAGACCCCAATTTTA 196
Qy 200 AAGGCCCAACCCGTGAACGTGACCTGCAACATCTTCAATCAACAGTTTTCAGTCCGTCAACA 259
Db 197 AAGNN 256
Qy 260 AGACCAATAGGACTACCGGGTGAATGTCTTTCTTGGGCAACAGTGGAAATGACCCACGCC 319
Db 257 NNN 316
Qy 320 TGTCTTACCGAGATATCTGTGACTCTCTGACCTCGATCCCTCCATGCTGGACTCTTA 379
Db 317 TCGGTACAGTGAATATCTGACGATCTTTAGACCTCGACCCCTCCATGTTGGACTCCA 376
Qy 380 TCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGCCAACTTCCATGAGGTGACCA 439
Db 377 TTTGGAACTGATTTGTTCTTCGCCAATGAAGAGGGTGCACACTTTCAATGAGTCACTA 436
Qy 440 CGGACAAAGTTPACTCGGCATCTTCAAGAATGGGAATGTGCTGTACACATCAGG---C 496
Db 437 CAGACAACAAAATGCTAAGAAATTTCAAAAAATGGAATGTCTTTATTCAATAAGGAGAT 496
Qy 497 TGNCCCTCATTTGCTCTGCTGATGACCTCAAGAACTTCCCATGGACATCCAGACGT 556
Db 497 TAACATTACACTTCTCTGCAATGGATCTCAAGAAATTTCCCATGGATGACAAACAT 556
Qy 557 GCACGATGAGCTGTGAGAGCTATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTT 616
Db 557 GTATATGCACTGGAAGCT----- 577
Qy 617 CAGTTGGCTACACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGTGTCTC 676
Db 578 ---TTGGGTACACAATGAATGATCTCATTTTGAATGGC---AAGATGAGGACCCCGTAC 631
Qy 677 AAGTGTGAGGGCTGACTCTGCCCACTTTATCTTGGGGATGAGAAGATCTAGGCT 736
Db 632 AAGTGTGAGGAGACTACACTTTGCCCGAGTTCTGTGTAAGAAAGAAAGATTTACGAT 691
Qy 737 GTGTGTACCAAGACTACACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCACTGTG 796
Db 692 ACTGCACTAACTTACATACAGGAAGTTTACGTGTATAGAGTGGCATTCATCTGG 751
Qy 797 AACGGCAGATGGGCTACTATCTGATTCAGATGTACATCCCGCCCTACTCATCTCATCC 856
Db 752 AGCGACAAATGGGATACTATCTGATCCAGATGTACATTCACAGTCTCTCTGATTGTTATTC 811
Qy 857 TGTCTTGGGTCTCTTCTGATCAACATGATGCTGCTCCCTGCGGCTGCGGCTGGCA 916
Db 812 TATCTTGGGTTCTGTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATG 871
Qy 917 TCACCACTGCTGTACATGACCAACAGAGCTCTGGCTCCCGGGCTCTTTTGCCTTAAGG 976
Db 872 TAACCACTGTCTAAAGATGACTACACAGATTCAGATTCAGATTCAGATTCAGATTCAG 931
Qy 977 TGTCTTACGTGAAGGCAATCGACATCTGAGTGGCTGTGTCTGTCTTGTGTGTCTGCTG 1036
Db 932 TTTTCATATGTCAAAGCTATTGATTTTGGATGGCAGTATGCTCTCTCTCTCTCTCTCTCAG 991
Qy 1037 CTTTGTCTGAGTATGCTGCCATAAATTTGTTTCTCTGTCAGCAATAAGAAATTCATAGCAG 1096
Db 992 CACTTCTGAGTATGACGTGTAAATTTTGTATCAAGACAAACAAGAACTTTNNNNNGAT 1051
Qy 1097 TTGGAAGAAGG-----CAGAGGCGCAACCTGAGGAAGATAT---CATCC 1141
Db 1052 TTCGACGAAGAGAAAGAAATACACAGAGCTTTTGCATTTGGAGAGTTTACCGTTTCT 1111
Qy 1142 AAGAAATCGTTTCTATTTCCGTGGCTATGGCTTGGGCCACTGCTGCGAGGCAAGAGATG 1201

Db 1112 CAGATATGATGATGAGGTAAAGGAAACCGGATTCAGCTTTCAGCCTATGGAATGGAC 1171
Qy 1202 GAGTCCCAATGAAGTCTTCTGCAAT-----TTATAGTCCCAACCTCCAGCCCC---TC 1252
Db 1172 CATGTCTCAAGCAAGAGATGGCATGCTCCAAAGGGCCCCCAACCTCTGTCAGGTAA 1231
Qy 1253 TTCTAAGGGAAGGAGAAACACCGCGAAACTCTAAGTGGACTGAGCAAGAAATTTGACA 1312
Db 1232 TGCCAAAAGTCTCTGATGAATGAGGAAGTCTTTATCGACCGGCCANNNNNNNGACA 1291
Qy 1313 CATCTCCCGGCTGTCTTCCCTTTTCACTTCTCTCATCTTCAATATCTTCTACTGGGTG 1372
Db 1292 CCATCTCCCGAGCTGCTTCCCATTTGTTTGAATTTTAATATTTTCTACTGGGTTA 1351
Qy 1373 TCTATAAAGTGTCTGCTGTCAGAAGATATCCA 1403
Db 1352 TCTATAAATTTCTTAGGCATGANNATATCA 1382

RESULT 12

CD354404
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

CD354404 773 bp mRNA linear EST 15-JUL-2003
UI-M-GMO-cgd-k-09-0-UI.r1 NIH_BMAP_GMO Mus musculus cDNA clone
IMAGE:30361304 5', mRNA sequence.
CD354404
CD354404.1 GI:31146905
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 773)
NIH-MGC <http://mgi.nhl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers
1. .773
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30361304"
/tissue_type="whole brain"
/dev_stage="1, 5 and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_GMO"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was size
selected according to mRNA size fraction, ligated with Ecor
I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CCAACTGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

FEATURES
source

ORIGIN


```

Query Match      27.0%; Score 442.4; DB 6; Length 773;
Best Local Similarity 74.2%; Pred. No. 4.5e-112;
Matches 607; Conservative 0; Mismatches 166; Indels 45; Gaps 2;

298 CAACAGTGAATACCCAGCCCTGTCTACCGAGATATCTCTGATGACTCTCTGGACCTC 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   1 CAGCAGTGGATGATTCACGGCTGGATACAGTGTAGTACCCAGATGATTCCTGGATTG 60

358 GATCCCTCAGTGTGACTCTATCTGGAAGCCAGACTCTCTCTTTGCTAATGAGAAAGG 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   61 GATCCCTCAATGTGGATTGATTCGATTGAAACCGGATTTGTTCTTGGCAATGAAAGG 120

418 GCCAATCTCCATGAGTGTACCGGACCGGACCAAGCTTACTGGCATCTTCAAGATGG 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   121 GCCAATTTCCATGATGTACACCTGACCAAGTGTGTGGGATTTCCAAAATGCAAA 180

478 GTCTCTGACAGCATCAGGCTGACCTCATTTGTCTGCTGCTGATGACCTCAAGAACTTC 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   181 GTCTCTCAGATTAAGACTCAGCTTGTACTTATCTGTGTCATGACCTGAAGAACTTT 240

538 CCATGAGCATCCAGACGTGACAGATGACAGTGTGAGAGCTCATCCATATCTTGCAGCCCT 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   241 CCAATGGATGTCAGACCTGTACAATGACGTGAGAGTT----- 280

598 CTGCCATCTCTGTCATCTTTCAGTGTGCTACACCATGAAGACCTCGTGTGAGTGCTG 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   281 -----TTGGGTACACCATGAATGACCTGATATTTGAGTGGTTA 318

658 GAAGATGCTCCTGCTGCCAAGTGGCTGAGGGGCTGACTCTGCCAGTTTATCTTGGCG 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   319 AGTATGGTC---CAGTACAAAGTTGCTGAAGAGTCACTCTTGCCCGAGTTTATTTGAAA 375

718 GATGAGAAGGATCTAGGCTGTTGTACCAAGCACTACAACACAGGGAATTCACCTGCATC 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   376 GAAGAGAGGAGCTTGGTTATTGSCAAGCAATTAACAACCTGGCAAGTTTACCTGCAIT 435

778 GAGTAAAGTTTCACTGGAAACGGCAGTGGCTTACTATCTGATTCAGATGTACATCCCC 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   436 GAGTCAAGTTTCACTGGAGCCAGATGGGCTACTATTGATCCAGATGTACATCCCC 495

838 AGCTACTCATCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   496 AGCTGTTGATGATCATTTTGTCTGGTCTCTCTTTTGGATAAATGAGATGAGCCCT 555

898 GCCCGTGTGGGCTGGGCTGACCAACCGTCTCACCATGACCAACCAAGCTCTGGCTCC 957
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   556 GCCAGGTTGCCCTGGGCTACCAACAGTCTGACATGACTACACAGATTGAGTTCC 615

958 CGGGCTCTTTGCTTAAGGTGCTCTACGTGAAGGCAATCGAATCTGGATGGTGTGT 1017
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   616 AGGGCATCTGCGCAAGGCTCTCTATGTGAAGCAATGACATCTGGATGGGCTATGC 675

1018 CTGCTCTTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   676 CTCTCTTTGTTGTTGCTGCTTACTGGAATATGACAGTGAATTTGTCTCAGGCAA 735

1078 CATAAAGAAATTCATAGACTTTCGAAGAGCAGAGCG 1115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   736 CATAAGGAGTTCTTGTCTCCGGAGACGACAGAG 773

```

```

RESULT 13
LOCUS      AU169868
DEFINITION AU169868 Ol-br-ad cDNA Oryzias latipes cDNA clone br5332, mRNA
sequence.
ACCESSION AU169868
VERSION   AU169868.1
KEYWORDS  EST.
SOURCE    Oryzias latipes (Japanese medaka)
ORGANISM  Oryzias latipes
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE
 AUTHORS Mita, K., Ishikawa, Y. and Yamauchi, M.
 TITLE Establishment of cDNA database of medaka, *Oryzias latipes*
 JOURNAL Unpublished (2001)
 COMMENT Contact: Mita K

Genome Research Group
 National Institute of Agrobiological Sciences
 Owaishi 1-2, Tsukuba, Ibaraki 305-8634, Japan
 Email: kmita@nias.affrc.go.jp
 method: uni-directional sequence direction: sequenced from T3 primer
 (5' -> 3').

FEATURES
 source Location/Qualifiers
 1..662
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="HNI"
 /db_xref="taxon:8090"
 /clone="br5332"
 /sex="female/male mixed"
 /tissue type="brain"
 /dev stage="adult"
 /clone_lib="Ol-br-ad cDNA"

ORIGIN

```

Query Match      24.5%; Score 401.8; DB 1; Length 662;
Best Local Similarity 76.4%; Pred. No. 9.9e-101;
Matches 540; Conservative 0; Mismatches 122; Indels 45; Gaps 2;

195 TTTTAAAGCCCAACCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGTCCGT 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   1 TTTTAAAGTTCACCTGTAACGTCACTGCAACATTTTATCAACAGCTTTGGATCCAT 60

255 CACCAAGACCAATGGACTACGGGTGAATGTCTTTCGGGCAACAGTGGAAATGACCC 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   61 CGCTGAACCAACATGGACTACAGATGAACATCTTCCTGAGGACAGAGTGGAAACGACC 120

315 AGCCTGTCTTACCAGAAATATCTGATGACTCTCTGGACCTCGATCCCTCCATGTCTGGA 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   121 CGGCTGGCTACAGCGAGTATCCGACGACTGCTGGATCTGGACCCGTCATGTGGA 180

375 CTCATCTGGAAGCCAGACCTCTCTTTGCTAATGAGAAAGGGGCAACTTCCATGAGGT 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   181 CTCATCTGGAAGCCGATCTGTCTTGTCTAATGAGAAAGGGGCAACTTCCACGAGGT 240

435 GACCAAGGACACAGTTACTCGGCATCTTCAAGATGGGAATGTCTGTACAGCATCAG 494
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   241 CACCAAGGACCAACAGCTGCTGCAATCTCCAAAATGGCAACGCTGTACAGCATACG 300

495 GCTGACCTCATTTTGTCTGCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAGAC 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   301 AATCACTCTGATCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

555 GTGACAGATGCAAGTTGAGAGCTATCCATATCTCTGACGCCCTCTGCGCATCTCTGTCACT 614
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   361 CTGCATCATGCAAGTGGAGAGCT----- 383

615 TTCAGTTGGCTACACCATGAAGACCTCGTTTGTGAGTGGCTGGAGATGCTCTCTGCTGT 674
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   384 -----TCGGCTACACCATGAACGACCTCATCTTCGAGTG---GGACGAGAAAGGGCCGT 435

675 CCAAGTGGCTGAGGGGCTGACTCTGCCCGAGTTTATCTTTCGGGATGAGAAAGATCTAGG 734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   436 GCAGTGGCGGAGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495

735 CTGTTGTACCAAGCACTACAACACAGGAAATTCACCTGCATCGAGGTAAAGTTTACCT 794
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   496 CTACTGCACCAAGCACTACAACACAGGTAATTTCACTGCTGATCGAGGCTCGCTTCCACCT 555

795 GGNACGGCAGATGGGCTACTATCTGATTGATGATGATGATGATGATGATGATGATGATGAT 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      556 GGAGGACAGATGGGCTACTACCTGATCCAGATGTATACATCCCTCGCTCATCGTCA 615
QY      855 CTTGCTCTGGGCTCTCTCTGATCAACATGATGCTGCGCTCGCC 901
Db      616 CTGTCTGGGCTCTCTCTGATCAACATGATGAGCGCGCGCGCC 662

RESULT 14
LOCUS   CD353729
DEFINITION UI-M-GMO-csc-m-22-0-UI.r1 NIH_BMAP_GMO Mus musculus cDNA clone
IMAGE:30360981 5', mRNA sequence.
CD353729
ACCESSION CD353729.1 GI:31146230
VERSION    EST.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus

REFERENCE 1 (bases 1 to 805)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Dr. Jim Lin, University of Iowa
          cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
          cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Distribution information can be found at
          http://genome.uiowa.edu/distribution/mousef1.html
          This clone was contributed by the Brain Molecular Anatomy Project
          (BMAP)

FEATURES
Seq primer: pyX-5.
Location/Qualifiers
1..805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30360981"
/tissue_type="whole brain"
/dev_stage="1, 5 and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GMO"
/notes="Organ: Brain; Vector: pyX-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was size
selected according to mRNA size fraction, ligated with Ecor
I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAACGGAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 24.2%; Score 397; DB 6; Length 805;
Best Local Similarity 71.9%; Pred. No. 2.3e-99;
Matches 536; Conservative 0; Mismatches 200; Indels 9; Gaps 1;

QY      663 TGTCTCTGCTGCTCAAGTGGCTGAGGGGCTGACTCTGCCCGGATTTATCTTGGGGATGA 722
Db      3 TGATGCTCCAGTACAAAGTTGCTGAGGAGCTCACCTTGCCCGGATTTATTTGAAAGAAGA 62

QY      723 GAAGGATCTAGGCTGTTGTACCAAGCACTACACACAGGGAATTCACCTGATCGAGGT 782
Db      63 GAAGGAGCTTGGTTATTGACAAAGCAATTACACACTGGCAAGTTTACCTGATTTGAGT 122

```

```

QY      783 AAAGTTTCACTGGAAACGCGCAGATGGGCTACTATCTGATTCAGATGTATATCCCGAGCT 842
Db      123 CAAAGTTTCACTGGAGCGCCAGATGGGCTACTATTTGATCCAGATGTATATCCCGAGCT 182
QY      843 ACTCATCGTCAATCCTGCTGGGCTCTCTCTGAGATCAACATGAGATGCTGCGCCG 902
Db      183 GTTGATAGTCAATTTTGTCTCTGGGCTCTCTCTTTGGATAAACATGATGCGAGCCCTGCCAG 242
QY      903 TGTGGGCTGGGCATCACACCGTCTCACCATGACCCAGAGCTCTGGCTCCCGGCG 962
Db      243 GGTTCGCCCTGGGCATCAACACAGTCTCTGCAATGACTACACAGAGTTCAAGGTTCCAGGGC 302
QY      963 CTCTTTGCTTAAGGTGTCTCTACGTGAAGCAATGCACATCTCGATGGCTGTGTCTGCT 1022
Db      303 ATCTCTGCCAAAGGTCTCTATGTGAAGCAATTCACATCTGATGGCGGTATGCTTCT 362
QY      1023 CTTTGTGTTGCTGCTCTGCTGAGATGTCTGCCATAAATTTTGTTCGTCAGCATAA 1082
Db      363 CTTTGTGTTGCTGCTCTGCTGAGATGTCTGCCAGTGAACCTTTGTCTCCAGGCAACATAA 422
QY      1083 AGAATTCATACACTTCGAAGAGCAGAGCGGCCAACCGCTTGGAGGAGATATCATCCA 1142
Db      423 GGAGTTCTCTGCTCTCCGAGACGACAGAGAGGAGCAATAAGAGAGAGATGTTACTCG 482
QY      1143 AGAAAGTGTCTTCTATTTCCGTGGCTATGGCTTGGGCCACTCTCGTCCAGGCAAGAGATGG 1202
Db      483 TGAAGTGTCTTAACTTCACTGAGTGGCTATGGGATGGTCACTGCTCCAAATGAAGATGG 542
QY      1203 AGGTCCAAATGGAAGTTCTGGCATTTATAGTCCCAACCTCCAGCCCTCTTTCTAAGGA 1262
Db      543 CACAGCTGTCAAGGTAC-----ACCTGCCAACCCACTTCCACAAACCCCAAGA 593
QY      1263 AGGAGAAACCGCGGAAACTCTAGTGGACTGAGCCAGAGAAATTCACCATCTCCCG 1322
Db      594 TGCAGATGCTATCAAGAAGAGTTTGTGGATCGGGGCAAAAGAAATTCACCATATCTCG 653
QY      1323 GGCTGTCTTCCCTTTTCACTTTCTCTCATCTTCAATATCTTCTACTGGGTTGTCTATAAAGT 1382
Db      654 AGCTGCTTCCCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 713
QY      1383 GCTATGGTCAGAAATATCCACCAG 1407
Db      714 CATTCGGCATGAAGATGCCACAAG 738

RESULT 15
CN266978
LOCUS     17000470616556 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN266978
ACCESSION CN266978.1 GI:47283392
VERSION    EST.
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 540)
AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Flak,G.J.,
Li,Y., Xu,C., Fang,R., Guegl,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE     Transcriptional characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT   Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 540 Std Error: 0.00.
Location/Qualifiers

```

```

source
1. .540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN
Query Match      23.1%; Score 378.6; DB 7; Length 540;
Best Local Similarity 99.0%; Pred. NO. 2.9e-94;
Matches 381; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 66 CCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAAATCTGGAACCAAGGGGTCCAGGCCAT 125
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 CCCAGGGGTGGCTTGGCAAAAGAGGAAGTCAAAATCTGGAACCAAGGGGGCCAGCCCAT 215
Qy 126 GTCCCTCTGATTTCTTAGACAACTTATGGGGCGAATCTGGATATGATGCCAGGAT 185
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 GTCCCTCTGATTTCTTAGACAACTTATGGGGCGAATCTGGATATGATGCCAGGAT 275
Qy 186 TCGGGCCCAATTTAAAGGCCACCCCTGACGTGACCTGCAACATCTTCATCAACAGTTT 245
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
276 TCGGGCCCAATTTAAAGGCCACCCCTGACGTGACCTGCAACATCTTCATCAACAGTTT 335
Qy 246 CAGCTCCGTACCAAGACCAATGACTACCGGGTGAATGCTCTTGGCGCAACAGTG 305
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
336 CAGCTCCGTACCAAGACCAATGACTACCGGGTGAATGCTCTTGGCGCAACAGTG 395
Qy 306 GAATGACCCAGCGCTCTCTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTC 365
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
396 GAATGACCCAGCGCTCTCTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTC 455
Qy 366 CATGCTGGACTCTATCTGAAGCCAGACCTCTTCTTGTATGAGAAAGGGGCCAATT 425
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
456 CATGCTGGACTCTATCTGAAGCCAGACCTCTTCTTGTATGAGAAAGGGGCCAATT 515
Qy 426 CCATGAGGTGACCAAGCAACAAG 450
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
516 CCATGAGGTGACCAAGCAACAAG 540

```

Search completed: April 28, 2005, 16:13:05
Job time : 6708.22 secs

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 39796)
 BURGESS, J.
 Direct Submission
 Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 humquerry@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Swi, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/ChrX
 LLOXNC01-35G3 is from the Lawrence Livermore National Laboratory flow-sorted X chromosome cosmid library LLOXNC01
 VECTOR: lawr16,
 Location/Qualifiers
 1..39796
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="RZPD:LLNCL110G0335"
 /db_xref="taxon:9606"
 /chromosome="X"
 /clone="LLOXNC01-35G3"
 /clone_lib="LLOXNC01"

 ORIGIN
 Query Match 91.7%; Score 44; DB 9; Length 39796;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 5 CATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTT 48
 |||||
 4077 CATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTT 4034

 RESULT 3
 AC137157
 LOCUS
 DEFINITION Mus musculus chromosome 8 clone RP23-377M11, WORKING DRAFT
 SEQUENCE, 9 unordered pieces.
 AC137157
 AC137157.2 GI:52077837
 HTG: HTGS PHAS1; HTGS DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (bases 1 to 194225)
 Wilson, R.K.
 The sequence of Mus musculus clone
 Unpublished
 2 (bases 1 to 194225)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (18-NOV-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 194225)
 Wilson, R.K.
 Direct Submission
 Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Sep 15, 2004 this sequence version replaced gi:25054283.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: M_BA0377M11
 ----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Chemistry: Dye-primer ET; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 190423 bases at least Q40
 Consensus quality: 190874 bases at least Q30
 Consensus quality: 191269 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 1386: contig of 1386 bp in length
 1387 1486: gap of unknown length
 1487 2690: contig of 1204 bp in length
 2691 2790: gap of unknown length
 2791 6363: contig of 3573 bp in length
 6364 18259: gap of unknown length
 18260 18359: contig of 11796 bp in length
 18360 35861: contig of 17502 bp in length
 35862 59175: contig of 23214 bp in length
 59176 59275: gap of unknown length
 59276 81565: contig of 22290 bp in length
 81566 127719: contig of unknown length
 127720 127820: contig of 46054 bp in length
 127821 194225: contig of unknown length
 194226 66406: contig of 66406 bp in length.
 Location/Qualifiers
 1..194225
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="8"
 /clone="RP23-377M11"
 1..1386
 /note="assembly_name:Contig14"
 1487..2690
 /note="assembly_name:Contig25"
 2791..6363
 /note="assembly_name:Contig26"
 6464..18259

FEATURES
 source
 misc_feature
 misc_feature
 misc_feature
 misc_feature

```

/note="assembly_name:Contig27"
18360 .35861
/note="assembly_name:Contig28
clone_end:17
vector_side:right"
35962 .59175
/note="assembly_name:Contig29"
59276 .81565
/note="assembly_name:Contig30"
81666 .127719
/note="assembly_name:Contig31"
127820 .194225
/note="assembly_name:Contig32
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 59.2%; Score 28.4; DB 2; Length 194225;
Best Local Similarity 76.1%; Pred. No. 3.1;
Matches 35; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CTATCATCTCTGCAGCCCTTGCACATCTGTCTGCTTCAGTT 48
|||||
Db 7117 CTCTCCAGACTCAGCAGGCTCTGCACATCTCTCTCTCTCAGTT 7162

RESULT 4
AC118255 200858 bp DNA linear ROD 09-SEP-2004
LOCUS Mus musculus chromosome 8, clone RP24-89A4, complete sequence.
DEFINITION
AC118255
ACCESSION
VERSION
KEYWORDS
SOURCE HTG.
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Birren,B., Nussbaum,C. and Lander,E.
1 (bases 1 to 200858)
MUS musculus chromosome 8, clone RP24-89A4, complete sequence.
TITLE
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 200858)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gordon,S., Goyette,M., Graham,L., Horton,L., Horton,L.,
Hagopian,D., Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Ribback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 200858)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cooke,P., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hago,B.,
Hagopian,D., Hago,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-SEP-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2004 this sequence version replaced gi:50345206.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@broad.mit.edu
----- Project Information
Center project name: L24331
Center clone name: 89_A_4
-----
Location/Qualifiers
1. .200858
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="8"
/map="8"
/clone="RP24-89A4"
/clone_lib="RP24-89A4"
1. .4
/note="clone boundary
clone end:SP6
site:MboI"
misc_feature
repeat_region
repeat_region

```

```
repeat_region /rpt_family="(CA)n"
968. .1018
repeat_region /rpt_family="(GA)n"
1410. .1674
repeat_region /rpt_family="B4A"
2267. .2593
repeat_region /rpt_family="IAPEY_LTR"
2614. .2708
repeat_region /rpt_family="(TG)n"
complement(2740. .2934)
repeat_region /rpt_family="B4"
3582. .3603
repeat_region /rpt_family="(CAG)n"
3961. .4094
repeat_region /rpt_family="B1F"
4507. .4569
repeat_region /rpt_family="CT-rich"
4577. .4754
repeat_region /rpt_family="(TTCC)n"
4912. .4941
repeat_region /rpt_family="(CAAAC)n"
4946. .5078
repeat_region /rpt_family="B1F"
5089. .5258
repeat_region /rpt_family="B3A"
6279. .6450
repeat_region /rpt_family="RSINE1"
6692. .6976
repeat_region /rpt_family="B4A"
complement(8257. .8393)
repeat_region /rpt_family="B1_MM"
complement(8424. .8549)
repeat_region /rpt_family="B1_MM"
8656. .8729
repeat_region /rpt_family="(TA)n"
complement(8748. .8952)
repeat_region /rpt_family="B4"
9711. .9784
repeat_region /rpt_family="(TTCC)n"
complement(9787. .9904)
repeat_region /rpt_family="PB1D10"
complement(10111. .10224)
repeat_region /rpt_family="PB1D9"
complement(10275. .10431)
repeat_region /rpt_family="RSINE1"
10894. .10944
repeat_region /rpt_family="(CA)n"
11519. .11602
repeat_region /rpt_family="(GA)n"
complement(11781. .11865)
repeat_region /rpt_family="B2L_S"
12746. .12771
repeat_region /rpt_family="(A)n"
complement(12956. .12965)
repeat_region /note="single clone coverage"
14874. .14902
repeat_region /rpt_family="(GGGA)n"
14987. .15030
repeat_region /rpt_family="(CAG)n"
15054. .15074
repeat_region /rpt_family="GC-rich"
15253. .15416
repeat_region /rpt_family="(CG)n"
15438. .15462
repeat_region /rpt_family="GC-rich"
16602. .16682
repeat_region /rpt_family="(TG)n"
17918. .17955
repeat_region /rpt_family="(TA)n"
18156. .18211
repeat_region /rpt_family="(T)n"
19441. .19483
repeat_region /rpt_family="(TTTTTG)n"
```

```
repeat_region 19689. .19719
/rpt_family="(CA)n"
repeat_region 20682. .20713
/rpt_family="AT-rich"
repeat_region 21432. .21452
/rpt_family="(TTTA)n"
repeat_region 22169. .22234
/rpt_family="ID4"
repeat_region complement(22530. .22666)
```

Query Match 59.2%; Score 28.4; DB 10; Length 200858;
Best Local Similarity 76.1%; Pred. No. 3.1;
Matches 35; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CTCATCCATACCTGCAGCCCTCTGCATCTCTGTGTCACATTCAGTT 48
|||||
Db 5229 CTCCTCCAGACTCAGCAGGCTCTGCATCTCTCTCTCTCAGTT 5274
|||||

RESULT 5
AC097996.2
WPCOMMENT
Sequence split into 8 fragments LOCUS AC097996 Accession AC097996
Fragment Name Begin End
AC097996.0 1 110000
AC097996.1 100001 210000
AC097996.2 200001 310000
AC097996.3 300001 410000
AC097996.4 400001 510000
AC097996.5 500001 610000
AC097996.6 600001 710000
AC097996.7 700001 711119

Continuation (3 of 8) of AC097996 from base 200001 (AC097996 Rattus norvegicus clone CH2;
Query Match 55.4%; Score 26.6; DB 2; Length 110000;
Best Local Similarity 78.0%; Pred. No. 15;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 TCATCCATACCTGCAGCCCTCTGCATCTCTGTGTCATTC 44
|||||
Db 77645 TCATCCATACCTCTTCAACCCCTCCACATCATCACCTTC 77685
|||||

RESULT 6
CR759777/c 175846 bp DNA linear HTG 04-SEP-2004
LOCUS
DEFINITION Danio rerio clone DKEY-162F5, WORKING DRAFT SEQUENCE, 8 unordered
pieces.
CR759777

ACCESSION
VERSION CR759777.2 GI:51890030
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM Danio rerio (zebrafish)

REFERENCE
AUTHORS Burton, J.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 4, 2004 this sequence version replaced gi:51889297.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK162F5
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads

Consensus quality: 173690 bases at least Q40
 Consensus quality: 174041 bases at least Q30
 Consensus quality: 174473 bases at least Q20
 Insert size: 175146; sum-of-contigs
 Quality coverage: 8.97x in Q20 bases; agarose-fp
 coverage: 8.44x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 2342: contig of 2342 bp in length
* 2343 2442: gap of 100 bp
* 2443 25852: contig of 23410 bp in length
* 25853 25922: gap of 100 bp
* 25923 25953 25954: contig of 71552 bp in length
* 25955 25956 25957: gap of 100 bp
* 25958 25959 25960: contig of 21383 bp in length
* 25961 25962 25963 25964: gap of 100 bp
* 25965 25966 25967 25968 25969 25970: contig of 31375 bp in length
* 25971 25972 25973 25974 25975 25976 25977 25978 25979 25980: contig of 6239 bp in length
* 25981 25982 25983 25984 25985 25986 25987 25988 25989 25990: gap of 100 bp
* 25991 25992 25993 25994 25995 25996 25997 25998 25999 26000: contig of 6150 bp in length
* 26001 26002 26003 26004 26005 26006 26007 26008 26009 26010: gap of 100 bp
* 26011 26012 26013 26014 26015 26016 26017 26018 26019 26020: contig of 12695 bp in length.

```

FEATURES

source

```

1. .175846
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /clone="DKFZ-162P5"
  /clone_lib="Daniokey"

```

misc_feature

```

1. .2342
  /note="assembly fragment:00027"

```

misc_feature

```

2443. .25852
  /note="assembly fragment:01017"

```

misc_feature

```

25953. .37504
  /note="assembly fragment:01383"

```

misc_feature

```

97605. .118987
  /note="assembly fragment:00369"

```

misc_feature

```

119088. .150462
  /note="assembly fragment:00670"

```

misc_feature

```

150563. .156801
  /note="assembly fragment:00059"

```

misc_feature

```

156902. .163051
  /note="assembly fragment:00104"

```

misc_feature

```

163152. .175846
  /note="assembly fragment:00182.0"

```

ORIGIN

Query Match 55.4%; Score 26.6; DB 2; Length 175846;
 Best Local Similarity 78.0%; Pred. No. 16;
 Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 CATCCATCTCTGACCCCTCTGCCATCTCTGTCACCTTCA 45
 |||||
 DB 170401 CATTCATAATCAGACGCTCTGCCATTTGTATCACTCTAA 170361

RESULT 7

AC129695/c

LOCUS

222983 bp DNA linear HTG 10-MAY-2003

DEFINITION

Rattus norvegicus clone CH230-11B16, WORKING DRAFT SEQUENCE, 2
 unordered pieces.

ACCESSION

AC129695

VERSION

GI:30522715

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 222983)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, S., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Geisregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshuwa, L., Louleghed, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokedemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
 Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G. and Gibbs, R. A.

TITLE

Direct Submission

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 222983)

AUTHORS

Worley, K. C.

TITLE

Direct Submission

JOURNAL

Submitted (01-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 222983)

AUTHORS

Rat Genome Sequencing Consortium.

TITLE

Direct Submission

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One

TITLE JOURNAL	Direct Submission Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 3 (bases 1 to 233746) Rat Genome Sequencing Consortium. Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	RESULT 9 AC094544/c LOCUS DEFINITION AC094544 VERSION AC094544.7 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)
REFERENCE AUTHORS TITLE JOURNAL	On May 10, 2003 this sequence version replaced gi:24819613. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)
COMMENT	Center: Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Project Information Center project name: GGPT Center clone name: CH230-171D15 Summary Statistics Assembly program: Atlas 3.0; Consensus quality: 223529 bases at least Q40 Consensus quality: 225655 bases at least Q30 Consensus quality: 227116 bases at least Q20 Estimated insert size: 232984; sum-of-contigs estimation Quality coverage: 8x in Q20 bases; sum-of-contigs estimation	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)
FEATURES source	* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. * 1 228528: contig of 228528 bp in length * 228529 228628: gap of unknown length * 228629 229841: contig of 1213 bp in length * 229842 229941: gap of unknown length * 229942 233746: contig of 3805 bp in length. Location/Qualifiers 1. 233746 /organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:10116" /clone="CH230-171D15" 1. 1039 /note="wgs_contig"	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)
ORIGIN	Query Match 55.4%; Score 26.6; DB 2; Length 233746; Best Local Similarity 78.0%; Pred No. 17; Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0; 4 TCATCCATCTCTGCGCCCTTCGCCATCTCTGTCACCTTC 44 Db 202426 TCATCCATCTCTCTCAACCCCTCCACATCTATCCTTC 202386 	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)	AC094544 Rattus norvegicus clone CH230-4118, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces. AC094544 GI:30466710 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED. Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus 1 (bases 1 to 233773)

TITLE
JOURNAL

COMMENT
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24942135.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GAWB
Center clone name: CH230-4118
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 229658 bases at least Q40
Consensus quality: 232789 bases at least Q30
Consensus quality: 235199 bases at least Q20
Estimated insert size: 238134; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 251012: contig of 251012 bp in length
* 251013 251112: gap of unknown length
* 251113 252217: contig of 1105 bp in length
* 252218 252317: gap of unknown length
* 252318 253973: contig of 1656 bp in length.

FEATURES
source
1. .253973
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-4118"
1. .1024
/note="wgs contig"
5428. .6613
/note="wgs contig"
36648. .37724
/note="wgs contig"
95090. .96935
/note="wgs contig"
245711. .246611
/note="clone_boundary
clone_end:17
site:EcoRI
end_sequence:BBH307674"
249761. .251012
/note="wgs end_extension
clone_end:17"

Query Match 55.4%; Score 26.6; DB 2; Length 253973;

ORIGIN

Best Local Similarity 78.0%; Pred. No. 17;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GCTCATCCATATCTGCGCCCTCTGCGCATCTCTGCACTT 42
Db 99678 GCTAGTCCAGAAATTACAGATATCTGCGCATCTCTGCACTT 99638

RESULT 10
AC123469/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-5003.6, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC123469
VERSION
AC123469.4 GI:25086546
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Chordata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE
1 (bases 1 to 259833)
AUTHORS
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, K., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louisedge, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., Von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
Direct Submission
Unpublished

```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 259833)
Worley, K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 259833)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23265401.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GXP2P
Center clone name: CH230-50016
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 241805 bases at least Q40
Consensus quality: 245039 bases at least Q30
Consensus quality: 247120 bases at least Q20
Estimated insert size: 247152; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 258704: contig of 258704 bp in length
* 258705 258804: gap of unknown length
* 258805 259833: contig of 1029 bp in length.
FEATURES
source
Location/Qualifiers
1. .259833
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-50016"
misc_feature
1. .1303
/notes="wgs end_extension
clone_end:Sp6"
6179. .6809
/notes="clone boundary
clone_end:Sp6
site:"
misc_feature
end sequence:BH336040"
complement(24927. .257242)
/notes="clone boundary
clone_end:T7
site:"
end sequence:BH336039"

```

```

misc_feature
68765. .70448
/note="wgs end_extension
clone_end:T7"
ORIGIN
Query Match 55.4%; Score 26.6; DB 2; Length 259833;
Best Local Similarity 78.0%; Pred. No. 17;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 4 TCATCCATCTCTGCAGCCCTCTGCATCTCTGTCACCTTC 44
|||||
DB 157894 TCATCCATCTCTTCAACCCCTCCACATCTACCTTC 157854
|||||
RESULT 11
AC137193 283158 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
*** 4 unordered pieces.
AC137193
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 283158)
Muzny, D. Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, B., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Stearle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K.,

```



```
Db 592 AGGTGCTGATCTCTGTTGCCTTCGCGCATGACTGTAAGTTTC 635
|||||
RESULT 13
AF348209/c 137436 bp DNA linear PRI 27-MAY-2001
LOCUS Homo sapiens IGF2R gene, complete cds.
DEFINITION AF348209.1 GI:14210065
ACCESSION AF348209.1
VERSION AF348209.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 137436)
Human M6P/IGF2R complete genomic sequence
Unpublished
2 (bases 1 to 137436)
Killian, J.K.
Direct Submission
Submitted (10-FEB-2001) Pathology, Duke University, USA
Location/Qualifiers
1. 137436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="6"
/map="6q26-q27"
join(2..296,22079..22218,39906..40030,41583..41681,
53471..55603,58084..58213,60449..60554,63450..63612,
63841..64006,65318..65421,71444..71608,74032..74172,
73398..75541,76629..76766,77382..77529,78043..78220,
78676..78791,79259..79427,81357..81536,87308..87409,
88907..89008,89790..89982,91431..91601,92387..92530,
92637..92812,93416..93503,94299..94514,95285..95415,
95688..95785,99133..99269,100752..100942,102794..102920,
103649..103768,104097..104353,104641..104859,
106731..106880,109085..109246,110465..110672,
111014..111160,114835..115069,115880..116016,
118918..119032,119992..120138,120801..120988,
127324..127510,133404..133556,134631..134700,
135559..137436)
/product="IGF2R"
join(148..296,22079..22218,39906..40030,41583..41681,
53471..55603,58084..58213,60449..60554,63450..63612,
63841..64006,65318..65421,71444..71608,74032..74172,
73398..75541,76629..76766,77382..77529,78043..78220,
78676..78791,79259..79427,81357..81536,87308..87409,
88907..89008,89790..89982,91431..91601,92387..92530,
92637..92812,93416..93503,94299..94514,95285..95415,
95688..95785,99133..99269,100752..100942,102794..102920,
103649..103768,104097..104353,104641..104859,
106731..106880,109085..109246,110465..110672,
111014..111160,114835..115069,115880..116016,
118918..119032,119992..120138,120801..120988,
127324..127510,133404..133556,134631..134700,
135559..137436)
/feature="IGF2R"
/codon_start=1
/product="IGF2R"
/protein_id="AAK56918.1"
/db_xref="GI:14210066"
/translation="MGAAGRSPHLGPARRPQRSLQLLLVAAPGSGTAQAAP
FPELCSYWEADVTKNNVLYKINICGSDIVQCGSSAVCMHDKTRTHVSGDSVLR
SATRSLEFNTVSCDQGGNHRVQSSIAFLCKGTGTEFVTECVHFWRITAA
CKKIDIFKANKVPCYVDFEELRKHDLNPLIKSGALYVDDSDPDSLSFINVCRDITL
RPGSQLACRACTPACRQAFQVQGRDLKLVKRDVLYSVREAGKGLDFCD
GHSPTVITFVCPSEERRETIPLKIAKSNCRVIEIWTYKADRVLESTCSLGEQ
QVSDILTPLAQSGLSYSDGKYLFLYNVCGETEIOFCNKKQAAVCKVKSQTSOV
KAAGRHNQTLRYSDGDLILIFGDECSGSGFQMSVINFECNKGAGNKGKTPFTG
EVDTYFTFTYETAYCKVEKEDLLCGATGKRYDLSALVRHAEPEQNEAVGSGQTE
```

TEKHHFFINICHRVLOEGKARGCPEDDAACVAVDKNGSKNLKFKFISSPMKEKGNOLSY
SGDDCGHGKKIKTNITLVCKFGDLESAPVLTITSGEGGCFYEFEMHTAAACVLSKTEG
ENCTVDFSOAGFSFDLSPKNGGAKVETKYDFYINVCSPVSPVSPVSPVSPVSPVSPV
KDEKTNLGNLSNAKLSYDGMILQNLRYGGTTPYNNERTHTPRATLITFLCDRAGVGP
EYDEEDNSTYNTFRWTSYACPEPELCTVTPSTLQYDLSSLAKEGGLGNGWYAMD
NGEHEVWRKYVYINVCPLNVPVCGNRYASAKMYKQSGQFTFVSTVSLNMAKMG
PVEEDSGSLLEYVNGSACTTSDGRTQTTTTRILHVCSPGRNLNHPDIFSLNMECVVP
LANTAAACPIQTITDQACIRDPNSGFFVNLNPLNSQQYVNSGIGIKFPHNVCGT
MPVCGTILGKPCSGCAEATQTEELKNKPARVPGIEKSLQSLSTEGTITLYKGLPSAK
GTADAFIVRFVSCDDVYSGPLFLHQIDISGQIRNTYEFETALACVPSVDCQVTD
LAGNMYDLTGLSTVRKPTAVDTSDVGRKRTFVLSVNCNRPITPGCGSAPVQLODG
KSNMLVVVQSPQAAAGSLSIMVNGDKCQNRSPRTIPECAOISGSAFQLODG
CYVIRWTEACPVVRVGDNCVKDPHGNLYDLKPLGNDTIVSAGEYTVYFVVC
GXLSDVCTFSDKSKVSCQKREPQGFHVKVAGLLTQKLYENGLLKNFTGGDTCG
KYQRTAIFFYCDRQTPVFLKSTCYLFEWRTYACPPFDTECSFGKDGKNS
FDLSLSRYSDNEAITGDPHEYLINVCYLAQAGTPECPPEAAACLLGSGKVN
LGRVDRGQWNGEIIIVKYVDGLDPCDGIKIKSTTIRPTCSQVNSRPMFISAVDC
EYTFWAPTACPMKSNHDDQVNTNPSHGLFDLSLSGRAGTAAAYSEKGLVYMSI
CGENENCPGVGACFCGQTRISVGKANKRLRYVDQVLYKDGSPCPSKSGLSYKSVI
SFVCPREAPTRNPMILSLDKQTCITLFSWHTPLACEQATECSVRNGSSIVDUSPLIH
RTGYEAYDESEDDASDNPDIYINICQPLNPMHGVPCPAGAAVCKVDPIDGPDIGR
VAGPPIPLNIAEYINFPESSTPCADKHFNTSLIAFHCXGVSMTGPKLRITSBD
FVPEWETPVVCPDEVMDGCTLTDEOLYSLNLSLSTSTFKVTRDSRYVSGVCTFA
VGPEOGCKGCGVCLLSGTGKASFGRLQSMKLDYRHQDRAVLSVYNGDRCPETDDG
VPCVPPFPGKSYEHCILIESRAKLCWSTADYDRDHENGFCRHSNSYRTSSIIPKCD
EDEDIGRPQVSEVRGCDVFEWTKVPCPPCKEYKQKHKTQDLRLSLGSGWS
LVHNGVSYINLCQIKYKGLPESASIRCTTTTGDVQVGLVHTKQLVGLGDKVW
TYSKGYPCGNKTASSVBLDTKTVPRAFPRFDIDCTTYTFSWDSRAACVAKPQEV
QMVNGTITNPKNGKSLGDIYFKLFRASGDMRTNGDNYLIEIQLSSIYSSNPACSG
ANICVQKNDQHFSRKSDTKYVLOQDGLDVVFASSSKCGKDKTKSVSTIFPHC
DPLVEDGIPFESHETADCOYLFSTWTSVAVCPGLGVGDFSDNPGDGMKHLGSRSSQV
GAVLSLLVALUCCLLALLYKKERRETVISKLTCCRRSSNVSYGKVKNEEETDE
NETEWMEIQLPPRQKGEQGHHTTKSVKALSSHLGDDQDSEDLVITPEVVKH
SGRGAASDSHPVRNAQNALQEREDDRVGLVRGEKARKGKSSSAQKTVSTKLVSF
HDDSDDELLHI"

ORIGIN

Query Match 55.0%; Score 26.4; DB 9; Length 137436;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGTCATCATCTCTGACGCTCTGCGCATCTCTGCTCACTTC 44
Db 43391 AGGTGCTGATCTCTGTTGCCTTCGCGCATGACTGTAAGTTTC 43348

AF069333 Homo sapiens insulin-like growth factor II receptor (IGF2R) gene,
partial cds. 138376 bp DNA linear PRI 17-SEP-2001

AF069333 AF069333 AF069334 AF069335 AF069336 AF069337 AF069338
AF069339 AF069340 AF069341 AF069342 AF069343 AF069344 AF069345
AF069346 AF069347 AF069348 AF069349 AF069350 AF069351 AF069352
AF069353 AF069354 AF069355 AF069356 AF069357 AF069358 AF069359
AF069360 AF069361 AF069362 AF069363 AF069364 AF069365 AF069366
AF069367 AF069368 AF069369 AF069370 AF069371 AF069372 AF069373
AF069374 AF069375 AF069376 AF069377 AF069378
AF069333.2 GI:15628184

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 138376)
Killian, J.K. and Jirtle, R.L.
Genomic structure of the human M6P/IGF2 receptor
Mamm. Genome 10 (1), 74-77 (1999)
99111400
PUBMED 9892739
2 (bases 1 to 138376)
Killian, J.K. and Jirtle, R.L.
Allelic variants human M6P/IGF2R

RESULT 14
AF069333/c
LOCUS AF069333
DEFINITION Homo sapiens insulin-like growth factor II receptor (IGF2R) gene,
partial cds.
ACCESSION AF069333 AF069334 AF069335 AF069336 AF069337 AF069338
AF069339 AF069340 AF069341 AF069342 AF069343 AF069344 AF069345
AF069346 AF069347 AF069348 AF069349 AF069350 AF069351 AF069352
AF069353 AF069354 AF069355 AF069356 AF069357 AF069358 AF069359
AF069360 AF069361 AF069362 AF069363 AF069364 AF069365 AF069366
AF069367 AF069368 AF069369 AF069370 AF069371 AF069372 AF069373
AF069374 AF069375 AF069376 AF069377 AF069378
AF069333.2 GI:15628184

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 138376)
Killian, J.K. and Jirtle, R.L.
Genomic structure of the human M6P/IGF2 receptor
Mamm. Genome 10 (1), 74-77 (1999)
99111400
PUBMED 9892739
2 (bases 1 to 138376)
Killian, J.K. and Jirtle, R.L.
Allelic variants human M6P/IGF2R

JOURNAL	Unpublished	3 (bases 1 to 138376)	On or before Sep 17, 2001 this sequence version replaced
REFERENCE	Killian, J.K. and Jirtle, R.L.		
AUTHORS	Direct Submission		
TITLE	Submitted (01-JUN-1998) Radiation Oncology and Pathology, Duke University Medical Center, Durham, NC 27710, USA		
JOURNAL	4 (bases 1 to 138376)		
REFERENCE	Killian, J.K. and Jirtle, R.L.		
AUTHORS	Direct Submission		
TITLE	Submitted (23-AUG-2000) Pathology, Duke University, Durham, NC 27710, USA		
JOURNAL	Sequence update by submitter		
REMARK	On or before Sep 17, 2001 this sequence version replaced		
COMMENT	gi:4165886, gi:3982689, gi:3982690, gi:3982691, gi:3982692, gi:3982693, gi:3982694, gi:3982695, gi:3982696, gi:3982697, gi:3982698, gi:3982699, gi:3982700, gi:3982701, gi:3982702, gi:3982703, gi:3982704, gi:3982705, gi:3982706, gi:3982707, gi:3982708, gi:3982709, gi:3982710, gi:3982711, gi:3982712, gi:3982713, gi:3982714, gi:3982715, gi:3982716, gi:3982717, gi:3982718, gi:3982719, gi:3982720, gi:3982721, gi:3982722, gi:3982723, gi:3982724, gi:3982725, gi:3982726, gi:3982727, gi:3982728, gi:3982729, gi:3982730, gi:3982731, gi:3982732, gi:3982733, gi:3982698.		
FEATURES	Location/Qualifiers		
source	1..138376		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosomes="6"		
	/map="6q26-27"		
	941..138376		
gene	/gene="IGF2R"		
mRNA	join(941..1236,23019..23158,40846..40970,42523..42621,56411..56543,59024..59153,61389..61494,64390..64552,64781..64946,66258..66361,72384..72548,74972..75112,76338..76481,77569..77706,78322..78469,78983..79160,79616..79731,80199..80367,82297..82476,88248..88349,89847..89948,90730..90922,92371..92541,93327..93470,93577..93752,94356..94443,95239..95454,96225..96355,96628..96725,100073..100209,101692..101882,103734..103860,104589..104708,105037..105293,105581..105799,107671..107820,10025..10186,111405..111612,111954..112100,115775..116009,116820..116956,119858..119972,120932..121078,121741..121928,128264..128450,134344..134496,135571..135640,136499..138376)		
CDS	/product="insulin-like growth factor II receptor"		
	join(1089..1236,23019..23158,40846..40970,42523..42621,56411..56543,59024..59153,61389..61494,64390..64552,64781..64946,66258..66361,72384..72548,74972..75112,76338..76481,77569..77706,78322..78469,78983..79160,79616..79731,80199..80367,82297..82476,88248..88349,89847..89948,90730..90922,92371..92541,93327..93470,93577..93752,94356..94443,95239..95454,96225..96355,96628..96725,100073..100209,101692..101882,103734..103860,104589..104708,105037..105293,105581..105799,107671..107820,110025..110186,111405..111612,111954..112100,115775..116009,116820..116956,119858..119972,120932..121078,121741..121928,128264..128450,134344..134496,135571..135640,136499..136909)		
	/gene="IGF2R"		
	/codon_start=1		
	/product="insulin-like growth factor II receptor"		
	/protein_id="AAL04402.1"		
	/db_xref="GI:15628185"		
	/translation="MCAAGRPHLPAPRRPQSLQLQLLLVAAPGSAQAAPFPFLCSTWEADVTKNVLKINIGSDVIVCGQSSAVCMHDLKTRVHSGDSVLRSAITSLLETNTVSCDQGNHRVQSSIAFLCGKILCTPFEVTAETCHYFHWRTAAKKDIFKANKEVPCVFDLELRHDLNPLIKUSGAYLVDDSDPDTSLFINVCRDITLLRQPSQLRACPGTGAALVRGQAFDVGQDRGLKRLKRLVLSYVREEAGKLDPCDGHSPAVTITFVCPSEERECTIPKLTAKSNCRYEIEWITEYACHRDYLESKTSLSGEQ		
ORIGIN	1 ACCTCATCATCTCTGCGAGCCCTCTGCATCTCTGCTCATCTTC 44		
Query Match	55.0%; Score 26.4; DB 9; Length 138376;		
Best Local Similarity	75.0%; Pred. No. 19;		
Matches	33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;		
AY293855	139947 bp DNA linear PRI 20-MAY-2003		
LOCUS	Homo sapiens insulin-like growth factor 2 receptor (IGF2R) gene, complete cds.		
DEFINITION	AY293855		
ACCESSION	AY293855		
VERSION	AY293855.1 GI:30962028		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 139947)		
AUTHORS	Rieder, M.J., Livingston, R.J., Daniels, M.R., Chung, M.-W., Miyamoto, K.E., Nguyen, C.P., Nguyen, D.A., Poel, C.L., Robertson, P.D., Schackwitz, W.S., Sherwood, J.K., Witrak, L.A. and Nickerson, D.A.		
TITLE	Submitted (07-MAY-2003) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA		
JOURNAL	To cite this work please use: NIH-SNPs, Environmental Genome Project, NIH-S E51478, Department of Genome Sciences, Seattle, WA		
COMMENT	(URL: http://egg.gs.washington.edu).		
FEATURES	Location/Qualifiers		
source	1..139947		


```

variation      /replace="g"
14292
/genes="IGF2R"
/frequency="0.02"
/replace="t"
14341
/genes="IGF2R"
/frequency="0.01"
/replace="t"
14462
/genes="IGF2R"
/frequency="0.04"
/replace="a"
14616
/genes="IGF2R"
/frequency="0.01"
/replace="a"
14655
/genes="IGF2R"
/frequency="0.24"
/replace="a"
14748
/genes="IGF2R"
/frequency="0.01"
/replace="t"
14887
/genes="IGF2R"
/frequency="0.21"
/replace="t"
14893..23111
/genes="IGF2R"
/notes="Region not scanned for variation"
misc_feature

```

```

Query Match      55.0%; Score 26.4; DB 9; Length 139947;
Best Local Similarity 75.0%; Pred.No. 19;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCTTTC 44
      |||||
Db      44630 AGGTGCTGCATCTCTGTGCTTCTGCCATGACTGTAGTTTC 44587

```

Search completed: April 28, 2005, 14:18:23
Job time : 234.469 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 01:49:33 ; Search time 31.5071 Seconds
(without alignments)
9018.523 Million cell updates/sec

Title: US-10-075-846-3_COPY_574_621

Perfect score: 48

Sequence: 1 agctcatcactactgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	1640	6 AAL49660	Aal49660 Human HGR
2	44	91.7	39796	3 AAC61681	Aac61681 Nucleotid
3	26.4	55.0	399	5 AAS85276	Aas85276 DNA encod
C 4	26	54.2	2848	4 AAS05331	Aas05331 Human alp
C 5	26	54.2	3135	4 AAS05316	Aas05316 Human alp
C 6	26	54.2	3322	4 AAS05322	Aas05322 Human alp
C 7	26	54.2	3558	4 AAS05317	Aas05317 Human alp
C 8	26	54.2	3745	4 AAS05323	Aas05323 Human alp
C 9	26	54.2	73063	12 ADQ97727	Adq97727 Human can
10	24	50.0	35100	2 AAV20441	Aav20441 Human c-f
11	24	50.0	38258	6 AAS98633	Aas98633 DNA encod
12	23.8	49.6	883	10 ADE11700	Adel11700 Human sec
13	23.8	49.6	49380	4 ABL11838	Abll11838 Drosophil
14	23.8	49.6	110000	12 ADQ59443_0	Adq59443 Human can
C 15	23.6	49.2	418	9 ACH39609	Ach39609 Human foe
16	23.6	49.2	33403	11 ACN44824	Acn44824 Mouse gen
17	23.6	49.2	46338	13 ABD33451	Abd33451 Murine ca
C 18	23.6	49.2	151858	13 ABD33489	Abd33489 Murine ca
19	23.4	48.8	423	4 AAK79953	Aak79953 Human imm
20	23.4	48.8	423	4 AAK79950	Aak79950 Human imm

C 21	23.4	48.8	2069	11 ADM01492	Adm01492 Human cDN
C 22	23.4	48.8	3717	13 ADR08147	Adr08147 Full leng
23	23.4	48.8	67832	9 ADA02801	Ada02801 Mouse itk
24	23.4	48.8	67832	10 ADB72539	Adb72539 Mouse itk
25	23.4	48.8	67832	10 ADC85281	Adc85281 Mouse itk
26	23.4	48.8	67832	12 ADM74396	Adm74396 Murine ca
27	23.2	48.3	454	9 ACH35812	Ach35812 Human end
C 28	23.2	48.3	2964	13 ACN37768	Acn37768 Tumour-as
29	23.2	48.3	5681	12 ADO19507	Ado19507 Human PRO
30	23.2	48.3	67093	13 ABD32671	Abd32671 Mouse can
C 31	23	47.9	717	13 ADS57091	Ads57091 Bacterial
C 32	23	47.9	3678	6 ABS66149	Abs66149 DNA encod
33	23	47.9	10386	4 ABL20052	Abll20052 Drosophil
34	23	47.9	16095	4 ABL04874	Abll04874 Drosophil
C 35	23	47.9	63886	11 ACN44158	Acn44158 Human gen
C 36	23	47.9	66315	12 ABQ76478	Abq76478 S. cerevi
C 37	22.8	47.5	1404	6 ABQ76478	Abq76478 S. cerevi
38	22.8	47.5	50013	11 ACN44708	Acn44708 Mouse gen
39	22.8	47.5	114931	11 ACN44422	Acn44422 Human gen
40	22.6	47.1	3876	6 ADI28076	Adi28076 ECMCAD ge
41	22.6	47.1	4667	6 ABA95199	Abas95199 Human car
42	22.6	47.1	6775	4 AAI58409	Aai58409 Human pol
43	22.6	47.1	6775	5 ADQ98619	Adq98619 DNA encod
44	22.6	47.1	6775	9 ADB48379	Adb48379 Novel hum
45	22.6	47.1	7008	4 AAI60195	Aai60195 Human pol

ALIGNMENTS

RESULT 1
AAL49660
ID AAL49660 standard; cDNA; 1640 BP.
XX
AC AAL49660;
XX
DT 27-NOV-2002 (first entry)
XX
DE Human HGRA4 splice variant coding sequence SEQ ID NO: 3.
XX
KW Human; glycine receptor alpha subunit 4; HGRA4; HGRA4sv; splice variant;
KW cardiovascular disorder; reproductive disorder; neural disorder;
KW cardiac; antiarrhythmic; antianginal; antiarrheic; antiulcer;
KW nocotropic; neuroprotective; antibacterial; virucide; protozoacide;
KW nervous system disorder; gastrointestinal disorder; gene therapy;
KW infection; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200266606-A2.
XX
PD 29-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-US0004329.
XX
PR 16-FEB-2001; 2001US-0269535P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemers N;
PI Chang H;
XX
XX WPI; 2002-674925/72.
DR P-PSDB; AAO19187.
XX
PT New isolated nucleic acid molecules encoding human glycine receptor A4
PT (HGRA4) polypeptides, useful for preventing, treating and ameliorating
PT conditions, e.g. neural or gastrointestinal disorders.
PS Claim 1; Fig 2; 349pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.

CC The sequences can be used in the treatment of neural disorders,
CC gastrointestinal disorders, disorders related to hyper glycine receptor
CC activity, cardiovascular disorders, reproductive disorders, or bacterial,
CC viral and parasitic infections. The present sequence is a coding sequence
CC of the invention

XX
SQ Sequence 1640 BP; 381 A; 456 C; 384 G; 419 T; 0 U; 0 Other;
Query Match 100.0%; Score 48; DB 6; Length 1640;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCATCTCTGCAGCCCTCTGCCATCTCTCTCACTTTCACTT 48
|||||
Db 574 AGCTCATCATCTCTGCAGCCCTCTGCCATCTCTCTCACTTTCACTT 621
|||||

RESULT 2
AAC61681
ID AAC61681 standard; DNA; 39796 BP.

XX AAC61681;

XX
DT 19-FEB-2001 (first entry)

DE Nucleotide sequence of the human ataxia gene.

XX Human; ataxia; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT exon 29850..29921

FT exon /*tag= a

FT intron 29922..33025

FT intron /*tag= b

FT exon 33026..33155

FT exon /*tag= c

FT intron 33156..33444

FT intron /*tag= d

FT exon 33445..33514

FT exon /*tag= e

FT intron 33513..33751

FT intron /*tag= f

FT exon 33752..33975

FT exon /*tag= g

FT intron 33976..34114

FT intron /*tag= h

FT exon 34115..34195

FT exon /*tag= i

FT intron 34196..35759

FT intron /*tag= j

FT exon 35760..35901

FT exon /*tag= k

FT intron 35902..38781

FT intron /*tag= l

FT exon 38782..38996

FT exon /*tag= m

XX WO200058461-A1.

XX 05-OCT-2000.

XX 23-MAR-2000; 2000WO-BP002600.

XX 26-MAR-1999; 99EP-00106343.

XX (RAPP/) RAPPOLD-HOERBRAND G.

XX Rappold-Hoerbrand G;

XX WPI; 2000-656166/63.

DR

XX

PT Novel nucleic acid sequence encoding human ataxia protein for screening
PT compounds useful for treating disorders relating to mutations in ataxia
PT gene.

XX Claim 6; Page 22-44; 47pp; English.

XX The present sequence represents the human ataxia gene. The ataxia protein
CC and polynucleotides are useful for diagnosing and treating disorders
CC related to ataxia. Ataxia gene sequences are useful in gene therapy, and
CC as diagnostic tools or reagents for identifying and characterizing
CC genetic defect involved in the disorders and diseases related to ataxia
XX

SQ Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 U; 0 Other;

Query Match 91.7%; Score 44; DB 3; Length 39796;

Best Local Similarity 100.0%; Pred. No. 6.3e-06;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CATCCATACCTCTGCAGCCCTCTGCCATCTCTCTCACTTTCACTT 48
|||||

Db 35720 CATCCATACCTCTGCAGCCCTCTGCCATCTCTCTCACTTTCACTT 35763
|||||

RESULT 3
AAS85276

ID AAS85276 standard; cDNA; 399 BP.

XX AAS85276;

XX
DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #21080.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX
PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639162/73.

XX P-PSDB; ABG21089.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 1; SEQ ID NO 21080; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 399 BP; 89 A; 110 C; 98 G; 102 T; 0 U; 0 Other;
Query Match 55.0%; Score 26.4; DB 5; Length 399;
Best Local Similarity 75.0%; Pred. No. 7.2;
Matches 33; Conservative 0; Mismatches 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTC 44
DB 35 AGGTCTCGATATCTGTGGCTTCTGCCATGACTGTAGTTC 78

RESULT 4
AAS05331/c
ID AAS05331 standard; DNA; 2848 BP.
XX
AC AAS05331;
XX
DT 07-SEP-2001 (first entry)
DE Human alpha-1,3 galactosyltransferase pseudogene genomic sequence #8.
XX
KW Human; alpha-1,3 galactosyltransferase promoter; livestock;
KW recombinant expression cassette; gene expression; transgenic animal;
KW growth hormone; xenotransplantation; transgenic cell; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Intron 1..65
FT /*tag= a
FT misc_feature 66..2676
FT /*tag= b
FT /*note= "Untranslated exon 9"
FT misc_feature 2677..2848
FT /*tag= c
FT /*note= "inter-gene sequence"
XX
PN WO200130992-A2.
XX
PD 03-MAY-2001.
XX
PF 20-OCT-2000; 2000WO-US029139.
XX
PR 22-OCT-1999; 99US-0161092P.
PR 25-AUG-2000; 2000US-0227951P.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Koike C;
XX
DR WPI; 2001-300505/31.
XX
PT New recombinant expression cassette useful in xenotransplantation.
PS Claim 16; Page 106-109; 138pp; English.
XX
CC The present sequence for human alpha-1,3 galactosyltransferase (GT) pseudogene genomic sequence #8 is 1 of 44 alpha-1,3 galactosyltransferase gene sequences (AAS05290-AAS05333) from different animals. The invention provides a novel recombinant expression cassette comprising an alpha-1,3 GT promoter operably linked to a polynucleotide for expression. The cassettes of the invention can be used to express foreign genes or to disrupt the native alpha-1,3 GT genomic sequence. The invention provides

CC transgenic animals and methods for their production. Transgenic mice in
CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can
CC be used for assessing promoter activity and specificity. Transgenic
CC livestock having the novel expression cassette in which a growth hormone
CC is expressed under the control of the alpha-1,3 GT promoter can be
CC matured or grown better than commonly employed strains. Tissue obtained
CC from transgenic animals can be implanted into a host providing a method
CC of xenotransplantation from a transgenic animal. The invention also
CC provides a transgenic organ consisting essentially of transgenic cells
CC engineered, for use in transplantation. The methods of the invention
CC facilitate xenotransplantation between species, particularly between
CC species exhibiting differential expression of the gal-alpha-gal epitope.
CC The invention also describes primers (AAS05334-AAS05385) used to isolate
CC the alpha-1,3 GT sequences
XX
SQ Sequence 2848 BP; 854 A; 558 C; 605 G; 826 T; 0 U; 5 Other;
Query Match 54.2%; Score 26; DB 4; Length 2848;
Best Local Similarity 76.2%; Pred. No. 15;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTC 44
DB 1727 CTCCTCAATCCACTGCAGGCCCTTGGCTTCTCTGACCACTC 1686

RESULT 5
AAS05316/c
ID AAS05316 standard; cDNA; 3135 BP.
XX
AC AAS05316;
XX
DT 07-SEP-2001 (first entry)
DE Human alpha-1,3 galactosyltransferase cDNA transcript sequence #1.
XX
KW Human; alpha-1,3 galactosyltransferase promoter; livestock;
KW recombinant expression cassette; gene expression; transgenic animal;
KW growth hormone; xenotransplantation; transgenic cell; ss.
XX
OS Homo sapiens.
XX
PN WO200130992-A2.
XX
PD 03-MAY-2001.
XX
PF 20-OCT-2000; 2000WO-US029139.
XX
PR 22-OCT-1999; 99US-0161092P.
PR 25-AUG-2000; 2000US-0227951P.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Koike C;
XX
DR WPI; 2001-300505/31.
XX
PT New recombinant expression cassette useful in xenotransplantation.
PS Claim 16; Page 80-83; 138pp; English.
XX
CC The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA transcript sequence #1 is 1 of 44 alpha-1,3 galactosyltransferase gene sequences (AAS05290-AAS05333) from different animals. The invention provides a novel recombinant expression cassette comprising an alpha-1,3 GT promoter operably linked to a polynucleotide for expression. The cassettes of the invention can be used to express foreign genes or to disrupt the native alpha-1,3 GT genomic sequence. The invention provides transgenic animals and methods for their production. Transgenic mice in which a reporter gene is operably linked to the alpha-1,3 GT promoter can be used for assessing promoter activity and specificity. Transgenic livestock having the novel expression cassette in which a growth hormone is expressed under the control of the alpha-1,3 GT promoter can be

CC matured or grown better than commonly employed strains. Tissue obtained
 CC from transgenic animals can be implanted into a host providing a method
 CC of xenotransplantation from a transgenic animal. The invention also
 CC provides a transgenic organ consisting essentially of transgenic cells
 CC engineered, for use in transplantation. The methods of the invention
 CC facilitate xenotransplantation between species, particularly between
 CC species exhibiting differential expression of the gal-alpha-gal epitope.
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate
 CC the alpha-1,3 GT sequences

XX Sequence 3135 BP; 945 A; 628 C; 704 G; 858 T; 0 U; 0 Other;

Query Match 54.2%; Score 26; DB 4; Length 3135;
 Best Local Similarity 76.2%; Pred. No. 15;
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 CTCATCCATCTCTGCAGCCCTTGCATCTCTGTCACCTTC 44
 |||||
 Db 2087 CTCCTCAATCCACTGCAGGCCCTTGCCTCTCTGACACCATC 2046
 |||||

RESULT 6
 AAS05322/c
 ID AAS05322 standard; cDNA; 3322 BP.

XX AAS05322;

DT 07-SEP-2001 (first entry)

XX Human alpha-1,3 galactosyltransferase cDNA transcript sequence #7.

XX Human; alpha-1,3 galactosyltransferase promoter; livestock;
 KW recombinant expression cassette; gene expression; transgenic animal;
 KW growth hormone; xenotransplantation; transgenic cell; ss.

XX Homo sapiens.

XX WO200130992-A2.

XX 03-MAY-2001.

XX 20-OCT-2000; 2000WO-US029139.

XX 22-OCT-1999; 99US-0161092P.

XX 25-AUG-2000; 2000US-0227951P.

XX (UYPI-) UNIV PITTSBURGH.

XX Koike C;

XX WPI; 2001-300505/31.

XX New recombinant expression cassette useful in xenotransplantation.

XX Claim 16; Page 94-97; 138pp; English.

XX The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA
 CC transcript sequence #7 is 1 of 44 alpha-1,3 galactosyltransferase gene
 CC sequences (AAS05290-AAS05333) from different animals. The invention
 CC provides a novel recombinant expression cassette comprising an alpha-1,3
 CC GT promoter operably linked to a polynucleotide for expression. The
 CC cassettes of the invention can be used to express foreign genes or to
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides
 CC transgenic animals and methods for their production. Transgenic mice in
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can
 CC be used for assessing promoter activity and specificity. Transgenic
 CC livestock having the novel expression cassette in which a growth hormone
 CC is expressed under the control of the alpha-1,3 GT promoter can be
 CC matured or grown better than commonly employed strains. Tissue obtained
 CC from transgenic animals can be implanted into a host providing a method
 CC of xenotransplantation from a transgenic animal. The invention also
 CC provides a transgenic organ consisting essentially of transgenic cells
 CC engineered, for use in transplantation. The methods of the invention

CC facilitate xenotransplantation between species, particularly between
 CC species exhibiting differential expression of the gal-alpha-gal epitope.
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate
 CC the alpha-1,3 GT sequences

XX Sequence 3322 BP; 991 A; 670 C; 756 G; 905 T; 0 U; 0 Other;

Query Match 54.2%; Score 26; DB 4; Length 3322;
 Best Local Similarity 76.2%; Pred. No. 16;
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 CTCATCCATCTCTGCAGCCCTTGCATCTCTGTCACCTTC 44
 |||||
 Db 2274 CTCCTCAATCCACTGCAGGCCCTTGCCTCTCTGACACCATC 2233
 |||||

RESULT 7
 AAS05317/c
 ID AAS05317 standard; cDNA; 3558 BP.

XX AAS05317;

XX 07-SEP-2001 (first entry)

XX Human alpha-1,3 galactosyltransferase cDNA transcript sequence #2.

XX Human; alpha-1,3 galactosyltransferase promoter; livestock;
 KW recombinant expression cassette; gene expression; transgenic animal;
 KW growth hormone; xenotransplantation; transgenic cell; ss.

XX Homo sapiens.

XX WO200130992-A2.

XX 03-MAY-2001.

XX 20-OCT-2000; 2000WO-US029139.

XX 22-OCT-1999; 99US-0161092P.

XX 25-AUG-2000; 2000US-0227951P.

XX (UYPI-) UNIV PITTSBURGH.

XX Koike C;

XX WPI; 2001-300505/31.

XX New recombinant expression cassette useful in xenotransplantation.

XX Claim 16; Page 83-86; 138pp; English.

XX The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA
 CC transcript sequence #2 is 1 of 44 alpha-1,3 galactosyltransferase gene
 CC sequences (AAS05290-AAS05333) from different animals. The invention
 CC provides a novel recombinant expression cassette comprising an alpha-1,3
 CC GT promoter operably linked to a polynucleotide for expression. The
 CC cassettes of the invention can be used to express foreign genes or to
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides
 CC transgenic animals and methods for their production. Transgenic mice in
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can
 CC be used for assessing promoter activity and specificity. Transgenic
 CC livestock having the novel expression cassette in which a growth hormone
 CC is expressed under the control of the alpha-1,3 GT promoter can be
 CC matured or grown better than commonly employed strains. Tissue obtained
 CC from transgenic animals can be implanted into a host providing a method
 CC of xenotransplantation from a transgenic animal. The invention also
 CC provides a transgenic organ consisting essentially of transgenic cells
 CC engineered, for use in transplantation. The methods of the invention
 CC facilitate xenotransplantation between species, particularly between
 CC species exhibiting differential expression of the gal-alpha-gal epitope.
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate
 CC the alpha-1,3 GT sequences

XX

```
Query Match      54.2%; Score 26; DB 4; Length 3745;
Best Local Similarity 76.2%; Pred. No. 16;
Matches 32: Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

XX Homo sapiens.

```
Query Match      54.2%; Score 26; DB 4; Length 3745;
Best Local Similarity 76.2%; Pred. No. 16;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```



```
PR XX 28-JUN-2001; 2001US-00892877.
PA (RUBE/) RUBEN S M.
PA (FLOR/) FLORENCE K A.
PA (NIJU/) NI J.
PA (ROSE/) ROSEN C A.
PA (CART/) CARTER K C.
PA (MOOR/) MOORE P A.
PA (OLSE/) OLSEN H S.
PA (SHIY/) SHI Y.
PA (YOUN/) YOUNG P E.
PA (WEIY/) WEI Y.
PA (BREW/) BREWER L A.
PA (SOPP/) SOPPET D R.
PA (LAFLE/) LAFLEUR D W.
PA (ENDR/) ENDRESS G A.
PA (EBNE/) EBNER R.
PA (BIRSE/) BIRSE C E.
XX Ruben SM, Florence KA, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei Y, Brewer LA, Soppet DR, Lafleur DW;
PI Endress GA, Ebner R, Birse CE;
XX WPI; 2003-801210/75.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX Claim 1; SEQ ID NO 72; 453pp; English.
XX The invention relates to human secreted polypeptides and the
CC polynucleotides encoding them. The sequences are useful for preparing
CC medicaments for preventing, treating or ameliorating medical conditions
CC e.g., cancer, liver disorders such as hepatitis or neural disorders such
CC as Alzheimer's disease. This sequence represents cDNA encoding a human
CC secreted polypeptide of the invention.
XX Sequence 883 BP; 141 A; 263 C; 251 G; 223 T; 0 U; 5 Other;
SQ
Query Match 49.6%; Score 23.8; DB 10; Length 883;
Best Local Similarity 80.0%; Pred. No. 76;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 TCATACCTCTGCAGCCCTCTGCCATCTCTGTCTACT 41
DB 701 TCATACCTCTGCAGCCGTATGCTTCCCTGCTACT 735
RESULT 13
ABL11838
ID ABL11838 standard; cDNA; 49380 BP.
XX
AC ABL11838;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 29996.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB67735.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 29996; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 49380 BP; 14230 A; 10576 C; 10238 G; 14336 T; 0 U; 0 Other;
SQ
Query Match 49.6%; Score 23.8; DB 4; Length 49380;
Best Local Similarity 72.1%; Pred. No. 1.8e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 AGCTCATCCATACCTCTGCAGCCCTCTGCCATCTCTGTCTACTTT 43
DB 16961 AGGTCTCCATATCTCGCGCCCTCCACAGTCTCTGTCTGTGT 17003
RESULT 14
ADQ59443.0
WP Sequence split into 4 fragments LOCUS ADQ59443 Accession Adq59443
WP Fragment Name Begin End
WP ADQ59443.0 1 110000
WP ADQ59443.1 100001 210000
WP ADQ59443.2 200001 310000
WP ADQ59443.3 300001 352938
ID ADQ59443 standard; DNA; 352938 BP.
XX
AC ADQ59443;
XX
XX 07-OCT-2004 (first entry)
XX
XX Human cancer-associated (CA) gene sequence SEQ ID NO:79.
XX
XX human; cancer-associated gene; cancer-associated protein; cytostatic;
KW gene therapy; vaccine; tyrosine kinase antagonist;
KW G-protein coupled receptor antagonist; cancer; lymphoma; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2004058288-A1.
XX
XX 15-JUL-2004.
XX
XX 15-DEC-2003; 2003WO-US040082.
XX
XX 17-DEC-2002; 2002US-00322696.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543349/52.
XX P-PSDB; ADQ59445.
XX
XX New cancer-associated nucleic acid for diagnosing, preventing or treating
```

PT cancer (e.g. lymphoma) or for screening agents that may be used for
PT treating or preventing cancer.
XX
PS Disclosure; SEQ ID NO 79; 143pp; English.
XX
CC The present invention describes human cancer-associated (CA) nucleotide
CC sequences (1). Also described: (1) an expression vector comprising (1);
CC (2) a host cell comprising (1) or the expression vector; (3) a microarray
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded
CC within an open reading frame of a CA sequence; (5) an isolated antibody,
CC or its antigen binding fragment, that binds to the above polypeptide; (6)
CC a hybridoma that produces the monoclonal antibody described above; (7) a
CC pharmaceutical composition comprising the antibody and a pharmaceutical
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising
CC the above (monoclonal) antibody or polynucleotide that selectively
CC hybridizes to any of the polynucleotide sequences mentioned above; (9)
CC methods for diagnosing cancer or for detecting the presence or absence of
CC cancer cells in an individual; (10) a method for inhibiting growth of
CC cancer cells in an individual; (11) a method for delivering a therapeutic
CC agent to cancer cells in an individual; (12) an electronic library
CC comprising the polynucleotide or polypeptide, or their fragments,
CC mentioned above; (13) a method of screening for anticancer activity; (14)
CC methods for detecting cancer associated with expression of a polypeptide
CC or the presence of the antibody in a test cell or serum sample; (15) a
CC method for screening for a bioactive agent capable of modulating the
CC activity of a CA protein encoded by the above nucleic acid molecule; and
CC (16) a method for treating cancers. (1) has cytostatic activity, and can
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,
CC and as a G-protein coupled receptor antagonist. The compositions and
CC methods of the present invention can be used for diagnosing, preventing
CC and treating cancer, especially lymphomas. They may also be used in
CC screening for agents that may be used for treating or preventing cancer.
CC The present sequence represents a human CA gene sequence, which is given
CC in the exemplification of the present invention. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 352938 BP; 94852 A; 72623 C; 75909 G; 104334 T; 0 U; 5220 Other;
Query Match 49.6%; Score 23.8; DB 12; Length 110000;
Best Local Similarity 72.1%; Pred. No. 2.1e+02;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 6 ATCCATCTCTGAGCCCTGCGCATCTCTGTCACCTTCAGTT 48
Db 46646 ATGCATTCTCTGAGACCACCTCCAGCCCTGTCGCAATTCACCT 46688
RESULT 15
ACH39609/c
ID ACH39609 standard; cDNA; 418 BP.
XX
AC ACH39609;
XX
XX 13-OCT-2003 (first entry)
XX
DE Human foetal brain cDNA #976.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW Genome mapping; biodiversity; genetic disorder.
OS Homo sapiens.
XX
XX US2003073623-A1.
PN
PD 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
PF
XX 30-JUL-2001; 2001US-00918995.
PR
XX (DRMA/) DRMANAC R T.
PA

PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 26821; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 418 BP; 145 A; 52 C; 129 G; 91 T; 0 U; 1 Other;
Query Match 49.2%; Score 23.6; DB 9; Length 418;
Best Local Similarity 86.7%; Pred. No. 78;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CATCCATCTCTGCGAGCCCTGCGCATCTC 34
Db 244 CCTCCAGACTCTGCGCATCCCTCTGCGCTTCTC 215
Search completed: April 28, 2005, 12:06:05
Job time : 38.5071 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 11:47:18 ; Search time 9.8673 Seconds
(without alignments)
7959.755 Million cell updates/sec

Title: US-10-075-846-3_COPY_574_621

Perfect score: 48

Sequence: 1 agctcatcactactctgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	50.0	35100	1	US-08-306-691B-19
2	24	50.0	35100	5	PCT-US93-06251-19
3	23.2	48.3	601	4	US-09-949-016-191370
4	23.2	48.3	42276	4	US-09-949-016-17218
5	23	47.9	601	4	US-09-949-016-135808
6	23	47.9	601	4	US-09-949-016-135925
7	23	47.9	601	4	US-09-949-016-136042
8	23	47.9	193555	4	US-09-949-016-15553
9	23	47.9	193555	4	US-09-949-016-15554
10	23	47.9	193555	4	US-09-949-016-15555
11	23	47.9	222452	4	US-09-949-016-12968
12	22.6	47.1	6775	4	US-09-949-016-12968
13	22.4	46.7	304533	4	US-09-949-016-15371
14	22.4	46.7	304533	4	US-09-949-016-15372
15	22.2	46.2	784931	4	US-09-949-016-15132
16	22	45.8	601	4	US-09-949-016-15132
17	22	45.8	825	1	US-08-248-474-109
18	22	45.8	825	3	US-08-756-849-109
19	22	45.8	44789	4	US-09-949-016-13909
20	21.8	45.4	601	4	US-09-949-016-127237
21	21.8	45.4	5761	3	US-09-323-472A-1
22	21.8	45.4	5761	3	US-09-323-472A-3
23	21.8	45.4	5761	3	US-09-323-472A-11
24	21.8	45.4	5761	4	US-09-585-077C-1
25	21.8	45.4	5761	4	US-09-585-077C-3
26	21.8	45.4	5761	4	US-09-585-077C-11
27	21.8	45.4	5761	4	US-09-585-077C-13

c 28	21.8	45.4	56939	4	US-09-949-016-13613	Sequence 13613, A
c 29	21.8	45.4	65561	4	US-09-949-016-15365	Sequence 15365, A
c 30	21.8	45.4	76281	4	US-09-949-016-12708	Sequence 12708, A
c 31	21.8	45.4	150780	4	US-09-949-016-14711	Sequence 14711, A
c 32	21.6	45.0	423	4	US-09-513-999C-28242	Sequence 28242, A
c 33	21.6	45.0	432	4	US-09-513-999C-1913	Sequence 1913, A
c 34	21.6	45.0	447	1	US-08-253-155A-22	Sequence 22, Appl
c 35	21.6	45.0	601	4	US-09-949-016-25301	Sequence 25301, A
c 36	21.6	45.0	601	4	US-09-949-016-25302	Sequence 25302, A
c 37	21.6	45.0	601	4	US-09-949-016-81829	Sequence 81829, A
c 38	21.6	45.0	601	4	US-09-949-016-81830	Sequence 81830, A
c 39	21.6	45.0	601	4	US-09-949-016-81831	Sequence 81831, A
c 40	21.6	45.0	601	4	US-09-949-016-81832	Sequence 81832, A
c 41	21.6	45.0	601	4	US-09-949-016-14260	Sequence 124260, A
c 42	21.6	45.0	601	4	US-09-949-016-14261	Sequence 124261, A
c 43	21.6	45.0	601	4	US-09-949-016-144834	Sequence 144834, A
c 44	21.6	45.0	601	4	US-09-949-016-144835	Sequence 144835, A
c 45	21.6	45.0	601	4	US-09-949-016-144836	Sequence 144836, A

ALIGNMENTS

RESULT 1
US-08-306-691B-19
; Sequence 19, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-19

Query Match 50.0%; Score 24; DB 1; Length 35100;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 GCTCATCTACTCTGCACCCCTCTGCCATCTCTGTCACT 41

Query Match 47.9%; Score 23; DB 4; Length 601;
Best Local Similarity 68.1%; Pred. No. 41;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Query Match 47.9%: Score 23: DB 4: Length 193555:

Best Local Similarity 68.1%; Pred. No. 1.5e+02;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 AGCTCATCATCTCTGCGCCCTCTGCCATCTCTGTCACTTTCACT 47
DB 73294 AGCTCTGCGCCCTCTGCACTGTGTGCGCCACCACTGCCACTGTGAAT 73248

RESULT 10
US-09-949-016-15555/c
; Sequence 15555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15555
; LENGTH: 193555
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(193555)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15555

Query Match 47.9%; Score 23; DB 4; Length 193555;
Best Local Similarity 68.1%; Pred. No. 1.5e+02;
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 AGCTCATCATCTCTGCGCCCTCTGCCATCTCTGTCACTTTCACT 47
DB 73294 AGCTCTGCGCCCTCTGCACTGTGTGCGCCACCACTGCCACTGTGAAT 73248

RESULT 11
US-09-949-016-12968
; Sequence 12968, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12968
; LENGTH: 222452
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12968
Query Match 47.9%; Score 23; DB 4; Length 222452;
Best Local Similarity 68.1%; Pred. No. 1.6e+02;

Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 2 GCTCATCATCATCTCTGCGCCCTCTGCCATCTCTGTCACTTTCACT 48
DB 67657 GCGCTCCCACTCTGCGCCCTCTGCCACACACATCTGAGATT 67703

RESULT 12
US-09-620-312D-289
; Sequence 289, Application US/09620312D
; Patent No. 6569682
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. 6569682el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 289
; LENGTH: 6775
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (197)..(4471)
US-09-620-312D-289

Query Match 47.1%; Score 22.6; DB 4; Length 6775;
Best Local Similarity 68.9%; Pred. No. 1e+02;
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 AGCTCATCATCTCTGCGCCCTCTGCCATCTCTGTCACTTTCA 45
DB 4032 AACCCATCCATTCAGCAGCCAGGTACAGCCTCTGCCACCTCCA 4076

RESULT 13
US-09-949-016-15371/c
; Sequence 15371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/241,755

Job time : 14.8673 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 08:01:55 ; Search time 89.6588 Seconds

(without alignments)
3259.123 Million cell updates/sec

Title: US-10-075-846-3_COPY_574_621

Perfect score: 48

Sequence: 1 agctcatcactactgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 11284434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCT05_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	1640	14	US-10-075-846-3
2	26.4	55.0	201	19	US-10-075-846-3
3	26.4	55.0	149382	19	US-10-075-846-3
4	26	54.2	2848	17	US-10-125-994A-42
5	26	54.2	3135	17	US-10-125-994A-27
6	26	54.2	3322	17	US-10-125-994A-33
7	26	54.2	3558	17	US-10-125-994A-28
8	26	54.2	3745	17	US-10-125-994A-34
9	24.2	50.4	548	18	US-10-437-963-43806
10	24	50.0	201	19	US-10-075-846-3
11	24	50.0	201	19	US-10-075-846-3

12	24	50.0	201	19	US-10-741-600-56505	Sequence 56505, A
13	24	50.0	72069	19	US-10-741-600-17775	Sequence 17775, A
14	23.8	49.6	883	10	US-09-892-877-73	Sequence 73, Appl
15	23.8	49.6	883	10	US-09-948-783-72	Sequence 72, Appl
16	23.8	49.6	352938	18	US-10-322-696-79	Sequence 79, Appl
17	23.6	49.2	263	17	US-10-424-599-31798	Sequence 31798, A
18	23.6	49.2	418	10	US-09-918-995-26821	Sequence 26821, A
19	23.6	49.2	33403	13	US-10-087-132-1465	Sequence 1465, Ap
20	23.6	49.2	46338	18	US-10-322-281-595	Sequence 595, App
21	23.6	49.2	151858	18	US-10-322-281-653	Sequence 653, App
22	23.4	48.8	608	13	US-10-027-632-187595	Sequence 187595,
23	23.4	48.8	608	17	US-10-027-632-187595	Sequence 187595,
24	23.4	48.8	2069	17	US-10-108-260A-177	Sequence 177, App
25	23.4	48.8	65300	19	US-10-741-600-17776	Sequence 17776, A
26	23.4	48.8	67832	11	US-09-997-722-67	Sequence 67, Appl
27	23.2	48.3	413	13	US-10-027-632-288605	Sequence 288605,
28	23.2	48.3	413	17	US-10-027-632-288605	Sequence 288605,
29	23.2	48.3	454	10	US-09-918-995-23024	Sequence 23024, A
30	23.2	48.3	67093	18	US-10-417-375-33	Sequence 33, Appl
31	23	47.9	655	13	US-10-027-632-12623	Sequence 12623, A
32	23	47.9	655	13	US-10-027-632-12624	Sequence 12624, A
33	23	47.9	655	13	US-10-027-632-12625	Sequence 12625, A
34	23	47.9	655	17	US-10-027-632-12623	Sequence 12623, A
35	23	47.9	655	17	US-10-027-632-12624	Sequence 12624, A
36	23	47.9	655	17	US-10-027-632-12625	Sequence 12625, A
37	23	47.9	717	17	US-10-369-493-32765	Sequence 32765, A
38	23	47.9	809	13	US-10-027-632-167429	Sequence 167429,
39	23	47.9	809	17	US-10-027-632-167429	Sequence 167429,
40	23	47.9	3678	17	US-10-432-443-140	Sequence 140, App
41	23	47.9	63686	13	US-10-087-192-466	Sequence 466, App
42	22.8	47.5	700	18	US-10-425-115-44389	Sequence 44389, A
43	22.8	47.5	1404	18	US-10-451-467A-381	Sequence 381, App
44	22.8	47.5	50013	13	US-10-087-192-1291	Sequence 1291, Ap
45	22.8	47.5	114931	13	US-10-087-192-862	Sequence 862, App

ALIGNMENTS

RESULT 1

US-10-075-846-3
; Sequence 3, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT EX
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075,846
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1293)
US-10-075-846-3

Query Match 100.0%; Score 48; DB 14; Length 1640;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCTATCCATCTCTGCGAGCCCTCTGCCATCTCTGTCTTTCAGTT 48

Db 574 AGTCTATCCATCTCTGCGAGCCCTCTGCCATCTCTGTCTTTCAGTT 621

RESULT 2

```
US-10-741-600-36664/c
; Sequence 36664, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36664
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-36664

Query Match          55.0%; Score 26.4; DB 19; Length 201;
Best Local Similarity 75.0%; Pred. No. 3.6;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 AGGTGCTCGATCTCTGTGCTTCTGCTTCTGCCATGACTGTAAGTTTC 151

RESULT 3
US-10-741-600-17661/c
; Sequence 17661, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17661
; LENGTH: 149382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(149382)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17661

Query Match          55.0%; Score 26.4; DB 19; Length 149382;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49389 AGGTGCTCGATCTCTGTGCTTCTGCTTCTGCCATGACTGTAAGTTTC 49346

RESULT 4
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-42/c

Query Match          55.0%; Score 26.4; DB 19; Length 149382;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 49389 AGGTGCTCGATCTCTGTGCTTCTGCTTCTGCCATGACTGTAAGTTTC 49346

RESULT 4
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-42/c
```

```
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-42/c

Query Match          54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1727 CTCCTCACTCACTGCAGCCCTTTCCTTCTCTGACCATC 1686

RESULT 5
US-10-125-994A-27/c
; Sequence 27, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-27/c

Query Match          54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1727 CTCCTCACTCACTGCAGCCCTTTCCTTCTCTGACCATC 1686

RESULT 6
US-10-125-994A-33/c
; Sequence 33, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-33/c

Query Match          54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2087 CTCCTCACTCACTGCAGCCCTTTCCTTCTCTGACCATC 2046

RESULT 6
US-10-125-994A-33/c
; Sequence 33, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-33/c
```

Query Match 50.0%; Score 24; DB 19; Length 201;
Best Local Similarity 75.0%; Pred. No. 28;

; PRIOR APPLICATION NUMBER: 60/231,846
 ; PRIOR FILING DATE: 2000-09-11
 ; PRIOR APPLICATION NUMBER: 09/892,877
 ; PRIOR FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: 09/437,658
 ; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/09847
 ; PRIOR FILING DATE: 1999-05-06
 ; PRIOR APPLICATION NUMBER: 60/085,093
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,094
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,105
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,180
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,927
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,906
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,924
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,922
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,921
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,923
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,925
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,928
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,920
 ; PRIOR FILING DATE: 1998-05-18
 ; NUMBER OF SEQ ID NOS: 465
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 72
 ; LENGTH: 883
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (8)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (28)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (30)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (47)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (873)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-948-783-72

Query Match 49.6%; Score 23.8; DB 10; Length 883;
 Best Local Similarity 80.0%; Pred. No. 36;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 TCCTACTCTGCAGCCCTCTGCCTACTCTGTCTACT 41
 |||||
 Db 701 TCTATACTCTCCAGCGGTATGTCTTCCCTGTCTACT 735

Search completed: April 28, 2005, 17:07:56
 Job time : 93.6588 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 03:05:59 ; Search time 195.782 Seconds

(without alignments)
9332.242 Million cell updates/sec

Title: US-10-075-846-3_COPY_574_621

Perfect score: 48

Sequence: 1 agctatccatctactgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	28.8	60.0	673	9	CE235158	CE235158 tigr-gss-
2	27.8	57.9	802	9	CL694721	CL694721 PRI0165a
3	26	54.2	231	4	BF948217	BF948217 CM2-NN115
4	26	54.2	235	7	CV101020	CV101020 UMC-pcrvb
5	26	54.2	264	7	CV101264	CV101264 UMC-pcrvb
6	26	54.2	284	7	CV101330	CV101330 UMC-pcrvb
7	26	54.2	265	7	CV101758	CV101758 UMC-pcrvb
8	25.8	53.8	665	8	BZ724707	BZ724707 OGEDS19TM
9	25.8	53.8	853	8	BZ724703	BZ724703 OGEDS19TC
10	25.4	52.9	1337	9	CG744715	CG744715 P037-2-F0
11	25.4	52.9	680	4	BI724543	BI724543 1031072F0
12	25.4	52.9	716	5	BU651556	BU651556 1112094C0
13	25.4	52.9	804	9	CW521717	CW521717 OP_Ba002
14	25	52.1	266	1	AL835518	AL835518 AL835518
15	25	52.1	358	1	AL842548	AL842548 AL842548
16	25	52.1	422	1	AL836072	AL836072 AL836072
17	25	52.1	426	1	AL835483	AL835483 AL835483
18	25	52.1	430	1	AI332005	AI332005 f966c12.y
19	25	52.1	441	1	AL835177	AL835177 AL835177
20	25	52.1	526	1	AL834659	AL834659 AL834659
21	25	52.1	547	8	AQ508245	AQ508245 RPCI-11-2
22	25	52.1	716	3	CR711148	CR711148 Tetraodon
23	25	52.1	806	7	CK467449	CK467449 938753 MA
24	25	52.1	889	7	CK465348	CK465348 936449 MA

25	24.8	51.7	490	6	CD739189	CD739189 4026644 1
c 26	24.8	51.7	610	8	AZ386786	AZ386786 IM0145N20
27	24.8	51.7	623	9	CR180890	CR180890 Forward s
c 28	24.8	51.7	675	9	CR064241	CR064241 Forward s
c 29	24.8	51.7	714	3	CR676321	CR676321 Tetraodon
c 30	24.8	51.7	860	9	CR002427	CR002427 Forward s
c 31	24.8	51.7	894	9	BX957628	BX957628 Forward s
c 32	24.8	51.7	970	9	CR179900	CR179900 Forward s
33	24.6	51.3	276	2	BB159552	BB159552 BB159552
34	24.6	51.3	402	9	CE052118	CE052118 tigr-gss-
35	24.6	51.3	454	7	CF602831	CF602831 BACCA01 0
36	24.6	51.3	586	7	CV093457	CV093457 FAMU USDA
37	24.6	51.3	658	7	CN547349	CN547349 EST 15338
38	24.6	51.3	670	7	CF215885	CF215885 CAST0002
39	24.6	51.3	710	6	CB885353	CB885353 OSJNEFISK
c 40	24.6	51.3	748	9	CC580162	CC580162 CH240 374
41	24.6	51.3	849	7	CF200536	CF200536 RR89015N0
42	24.6	51.3	897	7	CF213803	CF213803 CGF100080
c 43	24.4	50.8	329	7	CN236759	CN236759 WUB113B03
c 44	24.4	50.8	483	1	AU260810	AU260810 AU260810
45	24.4	50.8	584	7	CK943039	CK943039 4066948 B

ALIGNMENTS

RESULT 1
LOCUS CE235158/c 673 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-1700033378690 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION CE235158
VERSION CE235158.1 GI:35390916
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 673)
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
source
1. .673
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

Query Match 60.0%; Score 28.8; DB 9; Length 673;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 8 CCATCTCTGCAGCCCTCTGCCATCTCTGTCA 39
DB 32 CCACATCTGCAGCCCTCTGCCATCTCTGTCA 1

/note="Vector: GTGTAACCTA; Funding: A grant from the Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hepes and cultured in NCSU23 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were plunged into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NCSU23 for 12 or 24 hours. Library Construction (PCR protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml Rnase-free tube. Hybond messenger affinity paper was used to isolated the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the PCMV-SPORT6 vector. Preliminary library characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality (e.g. the presence of short polyA tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhulayan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG_TISSUE=Blastocyst after cryopreservation
TAG_SEQ=GTGTAACCTA"

ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 235;
Best Local Similarity 85.3%; Pred. No. 1.3e+02;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

13 CTCTGAGCCCTCGACCTCTGTCTTTCAG 46
|||||

Db 112 CTCTGAGCCCTCGACCTCTGTCTTTCAG 145

RESULT 5
CV101264
LOCUS
DEFINITION
264 bp mRNA linear EST 09-SEP-2004
scrofa cDNA 3', mRNA sequence.
CV101264
VERSION
CV101264.1 GI:51583428
KEYWORDS
EST.
SOURCE
Sus scrofa (pig)
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 264)
AUTHORS
Woods, R., Whitworth, K., Springer, G.K., Forrester, L.J., Spollen, W.G., Ries, J., Bivens, N., Murphy, C.N., Mathialigan, N., Agca, Y., Green, J.A. and Prather, R.S.
Swine Genomics
Unpublished (2004)
CONTACT: DNA Core Facility (Swine Project)
Animal Science - RS Prather
University of Missouri-Columbia
M616 Medical Sciences Bldg., Columbia, MO 65212, USA
Tel: (573)882-0428
Fax: (573)884-5552
Email: porcine@net.missouri.edu
POLYA=Yes.

FEATURES
Location/Qualifiers
1..264
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/dev_stage="Blastocyst after cryopreservation"
/clone_lib="pcrybst"
/note="Vector: GTGTAACCTA; Funding: A grant from the Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hepes and cultured in NCSU23 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were plunged into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NCSU23 for 12 or 24 hours. Library Construction (PCR protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml Rnase-free tube. Hybond messenger affinity paper was used to isolated the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the

PCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spolien, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG_TISSUE=Blasocyst after cryopreservation
TAG_SEQ=GTGTAACTTA

ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 264;
Best Local Similarity 85.3%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CTCTGAGCCCTCTGCCACTCTGTCTCATTTCAG 46
|||||
Db 111 CTCTGAGCCCTCTGCCACTCTGTCTCATTTCAG 144

RESULT 6

LOCUS CV101330 264 bp mRNA linear EST 09-SEP-2004
DEFINITION UMC-pcrybst1-008-d10 Blastocyst after cryopreservation pcryst Sus
scrofa cDNA 3', mRNA sequence.

ACCESSION CV101330

VERSION CV101330.1 GI:51583494

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 264)

Woods, R., Whitworth, K., Springer, G.K., Forrester, L.J.,

Spolien, W.G., Ries, J., Bivens, N., Murphy, C.N., Mathialigan, N.,

Agca, Y., Green, J.A. and Prather, R.S.

Swine Genomics

Unpublished (2004)

Contact: DNA Core Facility (Swine Project)

Animal Science - RS Prather

University of Missouri-Columbia

M616 Medical Sciences Bldg., Columbia, MO 65212, USA

Tel: (573)882-0428

Fax: (573)884-5552

Email: porcine@net.missouri.edu

POLYA=Yes.

Location/Qualifiers

1. .264

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/dev_stage="Blastocyst after cryopreservation"

/clone_lib="pcrybst"

/notes="Vector: GTGTAACTTA; Funding: A grant from the

ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 264;
Best Local Similarity 85.3%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CTCTGAGCCCTCTGCCACTCTGTCTCATTTCAG 46

|||||

Db 111 CTCTGAGCCCTCTGCCACTCTGTCTCATTTCAG 144

Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hapes and cultured in NCSU23 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were placed in into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NCSU23 for 12 or 24 hours. Library Construction (PCR Protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml kinase-free tube. Hybrid messenger affinity paper was used to isolated the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dr oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the pCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spolien, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG_TISSUE=Blasocyst after cryopreservation

TAG_SEQ=GTGTAACTTA

```

RESULT 7
CV101758
LOCUS
DEFINITION
  UMC-pcrybst1-014-f04 Blastocyst after cryopreservation pcrybst Sus
  scrofa cDNA 3', mRNA sequence.
ACCESSION
CV101758
VERSION
CV101758.1 GI:51583922
SOURCE
  Sus scrofa (pig)
  Sus scrofa
  Sus scrofa
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
  1 (bases 1 to 265)
  Woods, R., Whitworth, K., Springer, G.K., Forrester, L.J.,
  Spollen, W.G., Ries, J., Bivens, N., Murphy, C.N., Mathialigan, N.,
  Agca, Y., Green, J.A. and Prather, R.S.
  Swine Genomics
  Unpublished (2004)
  Contact: DNA Core Facility (Swine Project)
  Animal Science - RS Prather
  University of Missouri-Columbia
  M616 Medical Sciences Bldg., Columbia, MO 65212, USA
  Tel: (573)882-0428
  Fax: (573)884-5552
  Email: porcine@net.missouri.edu
  POLYA=Yes.

FEATURES
    source
        1..265
            Location/Qualifiers
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /dev_stage="Blastocyst after cryopreservation"
                /clone_lib="pcrybst"
                /note="Vector: GTGTAACTTA; Funding: A grant from the
                Monsanto Company to the University of Missouri. Genetic
                Source: Day 6 pig blastocyst stage embryos were flushed
                from a gilt's reproductive tract in TL Hepes and cultured
                in NCSU23 (1) for 3 hours prior to vitrification. For
                vitrification, embryos were cultured in 1M
                dimethylsulfoxide and 1.3M ethylene glycol in Tissue
                Culture Medium (TCM) and 20% newborn calf serum for 5
                minutes at 39 degrees Celsius. Then embryos were
                transferred to 3.2M ethylene glycol and 0.6M sucrose in
                TCM and 20% newborn calf serum for 45 to 60 seconds at 39
                degrees Celsius. Capillary action was used to load the
                embryos in open pulled straws and the straws were plunged
                into liquid nitrogen. At thawing embryos were placed in
                0.2M sucrose in TCM and 20% newborn calf serum for 1
                minute, then 0.1M sucrose in TCM and 20% newborn calf
                serum for 5 minutes. Embryos were then rinsed and cultured
                in NCSU23 for 12 or 24 hours. Library Construction (PCR
                Protocol): The amount of mRNA that was recovered from
                embryos was quite limiting and was not sufficient for
                library production with a standard protocol. Therefore,
                PCR-based protocol was utilized for producing libraries.
                The blastocysts were then rinsed with 50 ul
                polyvinylalcohol-PBS in depc treated medium three times
                and snap frozen with 1 ul super-RNasin in a 0.5ml
                RNase-free tube. Hybond messenger affinity paper was used
                to isolated the mRNA (2). The mRNA was reverse transcribed
                with a NotI-tag-dT18 oligonucleotide and a SMART
                oligonucleotide (Clontech) modified to contain a SalI site
                to generate full-length cDNA with a sequence complementary
                to the SMART oligonucleotide. Sequences within the SMART
                and dt oligonucleotides were used as primers to amplify
                the cDNAs by PCR with pfu turbo polymerase (Stratagene).
                The resulting PCR products were purified, digested with
                NotI and SalI and size fractionated by using a Chroma
                Spin-400 followed by a Chroma Spin-1000 column (Clontech).
                Purified cDNA from each PCR reaction was ligated into the
                pCMV-SPORT6 vector. Preliminary Library Characterization:

```

Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (-3 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG_TISSUE=Blastocyst after cryopreservation
TAG_SEQ=GTGTAACTTA"

ORIGIN

```

Query Match      54.2%; Score 26; DB 7; Length 265;
Best Local Similarity 85.3%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      13  CTCCTGAGCCCTCTGCCATCTCTGTCACCTTCAG 46
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      112  CTCCTGAGCCCTCTGCCATCTCTGTCACCTTCAG 145

```

RESULT 8

```

BZ724707/c
LOCUS      BZ724707
DEFINITION      665 bp DNA linear GSS 03-MAR-2003
      OGEDS19TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0253C13,
      genomic survey sequence.
ACCESSION      BZ724707
VERSION      BZ724707.1
KEYWORDS      GSS.
SOURCE      Zea mays
      ORGANISM
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
      clade; Panicoideae; Andropogoneae; Zea.

```

REFERENCE

```

AUTHORS      1 (bases 1 to 665)
      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
      Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
      Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
      Consortium for Maize Genomics
      Unpublished (2002)
      Other GSSs: OGEDS19TC
      Contact: Cathy Whitelaw
      TIGR
      9712 Medical Center Drive, Rockville, MD 20850, USA
      Tel: 301-838-5843
      Fax: 301-838-0208
      Email: whitelaw@tigr.org
      Seq primer: FR
      Class: sheared ends.
      Location/Qualifiers
          1..665
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone="ZMMBMA0253C13"

```

FEATURES

source

```

/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      53.8%; Score 25.8; DB 8; Length 665;
Best Local Similarity 73.3%; Pred. No. 1.9e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CTCATCCTACTCTGCAGCCCTCTGCACATCTCTGTCACTTCAGT 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 CTCATCCTACTTTCGGCGCGTGCACATCTCTATCAAAATTCAT 533

RESULT 9
BZ724703
LOCUS      853 bp      DNA      linear      GSS 03-MAR-2003
DEFINITION OGEDS19TC ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0253C13,
            genomic survey sequence.
ACCESSION  BZ724703
VERSION     BZ724703.1 GI:28697951
KEYWORDS   GSS.
SOURCE     Zea mays
            ORGANISM
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 853)
REFERENCE   1 (bases 1 to 853)
AUTHORS    Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
            Consortium for Maize Genomics
            Unpublished (2002)
JOURNAL
COMMENT     Other_GSSs: OGEDS19TM
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5943
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                1..853
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone_lib="ZM_0.7_1.5_KB"
                /clone="ZMMBMA0253C13"
                /notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
                methylation filtered genomic DNA library"

ORIGIN
Query Match      53.8%; Score 25.8; DB 8; Length 853;
Best Local Similarity 73.3%; Pred. No. 2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CTCATCCTACTCTGCAGCCCTCTGCACATCTCTGTCACTTCAGT 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 CTCATCCTACTTTCGGCGCGTGCACATCTCTATCAAAATTCAT 420

TIGR

RESULT 10
CG744715/c
LOCUS      1337 bp      DNA      linear      GSS 24-OCT-2003
DEFINITION P037-2-F02.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION  CG744715
VERSION     CG744715.1 GI:37965583
KEYWORDS   GSS.
SOURCE     Pristionchus pacificus
            ORGANISM
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
            1 (bases 1 to 1337)
REFERENCE   1 (bases 1 to 1337)
AUTHORS    Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
            An integrated physical and genetic map of the nematode Pristionchus
            pacificus
            Mol. Genet. Genomics 269 (5), 715-722 (2003)
            22835951
            12884007
            Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: raif.sommer@tuebingen.mpg.de
            Class: BAC ends.
            Location/Qualifiers
                1..1337
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Ppa EcoRI BAC Library"
                /note="The library was generated by a partial digest of
                the genomic DNA with EcoRI and cloning into the BAC
                vector."

ORIGIN
Query Match      53.8%; Score 25.8; DB 9; Length 1337;
Best Local Similarity 73.3%; Pred. No. 2.2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TCATCCTACTCTGCAGCCCTCTGCACATCTCTGTCACTTCAGT 48
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 851 TCATTCATCATCATCATCTCTCTCTCTCTCTCTCTCTCAATT 807

RESULT 11
BZ724543/c
LOCUS      680 bp      mRNA      linear      EST 19-SEP-2001
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION  BZ724543
VERSION     BZ724543.1 GI:15700238
KEYWORDS   EST.
SOURCE     Chlamydomonas reinhardtii
            ORGANISM
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaeae; Chlamydomonas.
            1 (bases 1 to 680)
REFERENCE   1 (bases 1 to 680)
AUTHORS    Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
            Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
            Unpublished (2001)
            Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
            Location/Qualifiers
                1..680
                /organism="Chlamydomonas reinhardtii"
                /mol_type="mRNA"
                /strain="CC-1690 wild type mt+ 21gr"
                /db_xref="taxon:3055"
                /clone_lib="C. reinhardtii CC-1690, Stress II
                (normalized), Lambda Zap II"

ORIGIN
Query Match      53.8%; Score 25.8; DB 8; Length 853;
Best Local Similarity 73.3%; Pred. No. 2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CTCATCCTACTCTGCAGCCCTCTGCACATCTCTGTCACTTCAGT 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 CTCATCCTACTTTCGGCGCGTGCACATCTCTATCAAAATTCAT 420

TIGR

RESULT 10
CG744715/c
LOCUS      1337 bp      DNA      linear      GSS 24-OCT-2003
DEFINITION P037-2-F02.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION  CG744715
VERSION     CG744715.1 GI:37965583
KEYWORDS   GSS.
SOURCE     Pristionchus pacificus
            ORGANISM
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
            1 (bases 1 to 1337)
REFERENCE   1 (bases 1 to 1337)
AUTHORS    Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
            An integrated physical and genetic map of the nematode Pristionchus
            pacificus
            Mol. Genet. Genomics 269 (5), 715-722 (2003)
            22835951
            12884007
            Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: raif.sommer@tuebingen.mpg.de
            Class: BAC ends.
            Location/Qualifiers
                1..1337
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Ppa EcoRI BAC Library"
                /note="The library was generated by a partial digest of
                the genomic DNA with EcoRI and cloning into the BAC
                vector."

ORIGIN
Query Match      53.8%; Score 25.8; DB 9; Length 1337;
Best Local Similarity 73.3%; Pred. No. 2.2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TCATCCTACTCTGCAGCCCTCTGCACATCTCTGTCACTTCAGT 48
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 851 TCATTCATCATCATCATCTCTCTCTCTCTCTCTCTCTCAATT 807

```

ORIGIN

Query Match 52.9%: Score 25.4: DB 5: Length 716:

Best Local Similarity	74.4%;	Pred. No. 2.7e+02;	
Matches	32: Conservative	0: Mismatches	11: Indels
		0: Gaps	0: Gaps

3

RESULT 13

DEFINITION OP_Ba0022J02.f OP_Ba Oryza punctata genomic clone OP_Ba0022J02 5', genomic survey sequence.

VERSION CW521717.1 GI:53995939

SOURCE

ORGANISM *Oryza punctata*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.

REFERENCE 1 (bases 1 to 804)
Ehrlartoideae; Oryzeae; Oryza.

Stum, D., Rao, K., Luo, M.

TITLE OMAP Project - Purdue University
JOURNAL Unpublished (2004)

Jackson Laboratory
Purdue University

915 W. State St., West Lafayette, IN 47907, USA
Tel: 7654963621

Fax: 7654967255
Email: sjackson@purdue.edu

Basecalling by phred version 0.020425.c. This sequence was derived from the raw sequence read by clipping with lucy version 1.19s.

Bases 42-845 of the raw sequence (length 1357) were retained after clipping.

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 161000 Std Error: 0.00

Plate: 0022 row: J column: 02
Seq primer: TAA TAC GAC TCA CTA TAG GG

Class: BAC ends.
Location/Qualifiers

```
source
1. .804
/organism="Oryza punctata"
```

```

/mol_type="genomic DNA"
/db_xref="taxon:4537"

```

```
/clone="OP_Ba0022J02"  
/tissue type="young leaves"
```

```
/lab host="DH10B-T1 phage resistant"  
/clone lib="OP Ba"
```

Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

Query Match 52.9%; Score 25.4; DB 9; Length 804;

Best Local Similarity 74.4%; Pred. No. 2.7e+02;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0

2 GCTCATCCATACTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44

[illegible]

Db 214 GCTCATCATCTCTCCGTCGGCCCTTTGCTGCACTGACATTTTC 256

RESULT 14
AL835518/c
LOCUS
DEFINITION
sequence.
ACCESSION
AL835518
VERSION
AL835518.1 GI:21877480
KEYWORDS
EST.
SOURCE
Takifugu rubripes (Fugu rubripes)
ORGANISM
Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
1 (bases 1 to 266)
Clark, M.S., Edwards, Y.J., Peterson, D., Clifton, S.W., Thompson, A.J.,
Sasaki, M., Suzuki, Y., Kikuchi, K., Watabe, S., Kawakami, K.,
Sugano, S., Elgar, G. and Johnson, S.L.
Fugu ESTs: new resources for transcription analysis and genome
annotation
Genome Res. 13 (12), 2747-2753 (2003)
Contact: Clark MS
MRC Human Genome Mapping Project Resource Centre
Hinxton, Cambridge, CB10 1SB, UK
Email: biohelp@hgmrc.ac.uk
Vector: pBlueScript II KS
V type: phagemid
PRIMER: KS

TITLE
The clone can be obtained from www.hgmrc.ac.uk

JOURNAL
COMMENT
Library created by Greg Elgar
MRC Human Genome Mapping Project Resource Centre
Cambridge, CB10 1SB, UK
Library sequenced by Sarah Warner and Jim Hills
MRC Human Genome Mapping Project Resource Centre, Hinxton,
Cambridge, CB10 1SB, UK

FEATURES
source
1..266
/location= "Takifugu rubripes"
/mol_type= "mRNA"
/db_xref= "taxon:31033"
/clone= "EPR025apck3"
/tissue_type= "spleen"
/clone_lib= "EPR0"
/note= "Vector: pBlueScript II KS"

ORIGIN
Query Match 52.1%; Score 25; DB 1; Length 266;
Best Local Similarity 75.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 3 CTCATCCATACCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
|||||
Db 167 CTCATCCAGAGTCAGCAGCTCTCTGGCAGCTTTGCGAATTT 127
|||||

RESULT 15
AL842548/c
LOCUS
DEFINITION
sequence.
ACCESSION
AL842548
VERSION
AL842548.1 GI:22020376
KEYWORDS
EST.
SOURCE
Takifugu rubripes (Fugu rubripes)
ORGANISM
Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
1 (bases 1 to 358)
Clark, M.S., Edwards, Y.J., Peterson, D., Clifton, S.W., Thompson, A.J.,
Sasaki, M., Suzuki, Y., Kikuchi, K., Watabe, S., Kawakami, K.,
Sugano, S., Elgar, G. and Johnson, S.L.
Fugu ESTs: new resources for transcription analysis and genome
annotation
Genome Res. 13 (12), 2747-2753 (2003)
Contact: Clark MS
MRC Human Genome Mapping Project Resource Centre
Hinxton, Cambridge, CB10 1SB, UK
Email: biohelp@hgmrc.ac.uk
Vector: pME18S-FL3
V type: phagemid
PRIMER: ME-735FW

TITLE
Library created by Koichi Kawakami, Masahide Sasaki, Yutaka Suzuki,
Sumio Sugano
The Institute of Medical Science, The University of Tokyo,
Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
and
Kiyoshi Kikuchi, Shugo Watabe
Laboratory of Aquatic Molecular Biology and Biotechnology, Graduate
School of Agricultural and Life Sciences, The University of Tokyo,
Bunkyo-ku, Tokyo 108-8639, Japan

JOURNAL
COMMENT
Library sequenced by Melody S. Clark and Amanda Thompson MRC Human
Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10
1SB, UK.

FEATURES
Location/Qualifiers
source
1..358
/organism= "Takifugu rubripes"
/mol_type= "mRNA"
/db_xref= "taxon:31033"
/clone= "F000H06aF1"
/sex= "female"
/tissue_type= "heart"
/dev_stage= "adult"
/clone_lib= "F000H"
/note= "Vector: pME18S-FL3"

ORIGIN
Query Match 52.1%; Score 25; DB 1; Length 358;
Best Local Similarity 75.6%; Pred. No. 3.3e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 3 CTCATCCATACCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
|||||
Db 199 CTCATCCAGAGTCAGCAGCTCTCTGGCAGCTTTGCGAATTT 159
|||||

Search completed: April 28, 2005, 16:13:12
Job time : 202.782 secs